MASTER'S DEGREE THESIS

Master of Science in Advanced Mathematics and Mathematical Engineering

Title: Congruences Between Modular Forms

Author: Oriol Fernández Peña

Advisor: Dr. Víctor Rotger Cerdà

Department: Departament de Matemàtiques

Academic year: 2017-2018

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Universitat Politècnica de Catalunya

Facultat de Matemàtiques i Estadística

Master in Advanced Mathematics and Mathematical Engineering

Master's Thesis

Congruences between modular forms

Oriol Fernández Peña

Supervisor: Dr. Víctor Rotger

June, 2018

Abstract

This master's thesis is intended to give a presentation of the theory of congruences between the Fourier coefficients of modular forms. In order to do that we introduce the reader to the basic theory of modular forms from the beginning and we study the structure of their Fourier coefficients in different ways using Hecke operators. Then we start the theory of congruences finding some of them by classical methods of Number Theory. After that, we introduce the advances made by Swinnerton-Dyer in the study of congruences using ℓ-adic representations and the generalisation by Ken Ono. Finally, we explain the papers by Hida and Ribet in two chapters giving some conditions for the existence of congruences using the associated L-functions and decomposing the space of modular forms.

Resum

Aquest treball final de màster té com a objectiu fer una presentació de la teoria de congruències entre els coeficients de Fourier de formes modulars. Per a fer això introduïm al lector a la teoria bàsica de formes modulars des del principi i estudiem l'estructura dels seus coeficients de Fourier mitjançant els operadors de Hecke. Després, encetem la teoria de congruències estudiant algunes d'elles per mitjans clàssics de la Teoria de Nombres. Un cop introduït el concepte de congruència expliquem els avenços de Swinnerton-Dyer que va fer servir representacions ℓ-àdiques per trobar congruències i les generalitzacions de Ken Ono. Finalment, expliquem les publicacions de Hida i Ribet en dos capítols i donem algunes condicions per l'existència de congruències fent servir la L-funció i descomponent l'espai de formes modulars.

Keywords

Number Theory, Modular Forms, Congruences between modular forms.

Contents

Introduction

The study of modular forms began in the first years of the nineteenth century in the works of Jacobi treated as theta functions. It appears too in his works about binary quadratic forms.

When Riemann started the systematic study of some surfaces with good properties, what we call now in his honour Riemann surfaces, it was soon popularised between geometers the study of the quotients of the Poincaré half plane by the action of some subgroups of $\mathrm{SL}_2(\mathbb{Z})$, the so-called modular curves.

It seems that it was Klein who first used the word *modulform* in [FK90], however it was Hecke who formalised the concept of modular forms and developed the theory of what he called averaging operators, and now we know by Hecke operators. Apparently, the first time the word *modulform* appears in his work is in [Hec24].

Hecke was crucial in the development of the theory of modular forms, when he defined the Hecke operators, he soon realised that the Fourier coefficients of the eigenforms had arithmetical properties, not only that they lie in some number field, which is extremely important if we want to do Number Theory, but that they could define an L-function which had an Euler product

$$L(f,s) = \sum_{n\geq 1} \frac{a_n}{n^s} = \prod_p (1 - a_p p^{-s} + \chi(p) p^{k-1-2s})^{-1},$$

where $\chi : \mathbb{Z}/N\mathbb{Z} \longrightarrow \mathbb{C}$ is a Dirichlet character and $f = \sum_{n\geq 1} a_n q^n \in \mathcal{S}_k(N,\chi)$.

In this thesis we will benefit from these good properties that Hecke found, our main goal is study from different points of view how we can find congruences between the Fourier coefficients of modular forms. For this we need to develop all the theory of modular forms.

Our question tries to understand what is the underlying structure of the spaces $\mathcal{M}_k(\Gamma)$ and $\mathcal{S}_k(\Gamma)$ modulo primes. It seems that it was Ramanujan the first one who proposed a congruence

$$\tau(n) \equiv \sum_{d|n} d^{11} \pmod{691},$$

where $\tau(n)$ called the τ of Ramanujan is the n-th coefficient of

$$\Delta(z) = q\prod_{n\geq 1}(1-q^n)^{24}, \quad q = e^{2\pi iz}.$$

We will see later that this is the same as

$$\Delta \equiv G_{12}(z) \pmod{691},$$

where $G_k(z)$ is the Eisenstein series of weight k.

This congruence can be proved in many ways and we will see them all along the thesis. Soon after, Swinnerton-Dyer made a discovery, he used ℓ-adic representations attached to modular forms and studied their images in order to find the structure of the Fourier coefficients modulo ℓ. This is a immense result since it gives us all congruences, the bad point about the theory of Swinnerton-Dyer is that it only covers totally the case where these Fourier coefficients are rational numbers. However, Ono generalised these results to a more general setting.

On the other hand, the treatment Hida gave was analytical and uses some special values of the L-function associated to the modular form. Very similar to this approach Ribet studied some reciprocal questions.

All these approaches made important advances in the problem, however, all the results throughout the thesis are either computational or existencialist. In other words, either we need to suspect that two modular forms are congruent modulo a prime p and then we have the tools to prove it or disprove it or we can assure the existence of a modular form congruent to the one we study modulo some prime. The methods to find them are almost always *ad hoc* and the problem of finding all congruences modulo a prime is still open.

Structure of the thesis and prerequisites

The thesis is divided into six chapters. Each one tries to introduce different points of view in order to solve the problem we put. The structure is rather chronological, with some exceptions, and introduces the results in the way they were published.

- The first chapter treats the basics of the theory of modular forms, we introduce the concepts of modular form, cusp form, nebentype, level and we talk about modular forms modulo primes. We give some examples and describe the structure of these vector spaces. We follow essentially the texts of [DS05] and [Miy06]. There is no special prerequisites, but we use complex analysis together with Fourier theory and finite fields.

- The second chapter introduces the concepts of Hecke operators and the Petersson inner product. We prove that there is always a basis of Hecke forms and that their Fourier coefficients lie in some number field which justifies the goal of our thesis.
- In the third chapter we start the treatment of the Fourier coefficients in an algebraic way, we study and present some results which use nothing but basic facts on Number Theory. Even if the name say classical congruences, the fact is that not all results are classical but all of them use classical methods. We arrive to the famous Ramanujan's conjecture by different means.
- In this chapter we study ℓ-adic representations and the results by Swinnerton-Dyer and Ono, which using the images of the attached ℓ-adic representation gave congruences on the Fourier coefficients.
- This chapter gives an sketch of the proof of a strong theorem by Hida. There are some gaps which on parabolic cohomology and étale cohomology, but these are not important in order to understand the core of the proof.
- The last chapter tries to give an insight of the results of Ribet and proves the existence of congruences between oldforms and newforms.

Acknowledgements

I feel very grateful to my advisor Dr. Víctor Rotger for guiding me during this journey discovering new topics in Number Theory and for helping me out to understand those hard and easy proofs I did not follow.

I am also very grateful to my family and my partner for supporting me on the good and on the bad moments.

Chapter 1

Modular forms

The aim of this chapter is introduce the reader to the basics of the theory of modular forms. We need to understand well all this theory in order to understand the rest of the thesis. We follow basically the texts [DS05] and [Miy06].

1.1 Definitions

In all this thesis $\mathbb{H}$ will denote the upper half plane of $\mathbb{C}$, $\mathcal{H}(U)$ the holomorphic functions of $U \subseteq \mathbb{C}$. $\mathrm{SL}_2(\mathbb{Z})$ will be the group of 2×2 matrices with integral entries and determinant equal to 1 and $\mathrm{GL}_2(\mathbb{Q})^+$ the rational matrices with positive determinant.

Definition 1.1.1. *Let $N \in \mathbb{N}$. The **principal congruence subgroup of level** N is*

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod N \right\}.$$

*A **congruence subgroup of level** N is any subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma(N) \subseteq \Gamma$.*

Examples 1.1.2. *Two important examples are the following. For $N \in \mathbb{N}$*

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod N \right\}.$$

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod N \right\}.$$

Proposition 1.1.3. *For any integer $N \geq 0$*

$$\Gamma_1(N)/\Gamma(N) \simeq \mathbb{Z}/N\mathbb{Z},$$

$$\Gamma_0(N)/\Gamma_1(N) \simeq (\mathbb{Z}/N\mathbb{Z})^*.$$

Proof. This is easy, define

$$\begin{array}{ccc} \Gamma_1(N) & \longrightarrow & \mathbb{Z}/N\mathbb{Z} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \longmapsto & b. \end{array}$$

It is a morphism because modulo N

$$\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & b' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & b+b' \\ 0 & 1 \end{pmatrix}.$$

It is clearly surjective. The kernel is

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod N \right\} = \Gamma(N).$$

The other one is very similar, define the following morphism

$$\begin{array}{ccc} \Gamma_0(N) & \longrightarrow & (\mathbb{Z}/N\mathbb{Z})^* \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \longmapsto & d. \end{array}$$

it is a morphism because $c \equiv 0 \pmod N$, so the d entry behaves multiplicatively. Moreover, it is surjective: let $d \in \mathbb{Z}$ coprime to N, then there is a such that $ad \equiv 1 \pmod N$ i.e., there is $r \in \mathbb{Z}$ such that $ad = 1 + Nr$, choose the matrix

$$\begin{pmatrix} a & r \\ N & d \end{pmatrix}.$$

Now, the kernel are the matrices such that $d \equiv 1 \pmod N$, hence since the determinant modulo N is $ad \equiv 1 \pmod N$ we have $a \equiv 1 \pmod N$, i.e., the matrix is in $\Gamma_1(N)$. The other inclusion is trivial. □

Definition 1.1.4. *For any $\gamma \in \mathrm{GL}_2(\mathbb{Q})^+$ and $f : \mathbb{H} \longrightarrow \mathbb{C}$ function let*

$$f|[\gamma]_k(z) = \det(\gamma)^{k-1}(cz+d)^{-k} f(\gamma(z)),$$

where $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *and* $\gamma(z) = \frac{az+b}{cz+d}$.

Where $\mathrm{GL}_2(\mathbb{Q})^+ = \{\gamma \in \mathrm{GL}_2(\mathbb{Q}) : \quad \det(\gamma) > 0\}$.

Definition 1.1.5. *Let f be a function from $\mathbb{H}$ to $\mathbb{C}$, $k \in \mathbb{Z}$ and Γ a congruence subgroup. We say that f is a **modular form of weight** k if the following conditions hold:*

1. $f \in \mathcal{H}(\mathbb{H})$.

2. *For any* $\alpha \in \Gamma$, $f|[\alpha]_k(z) = f(z), \quad \forall z \in \mathbb{H}$.

3. f *is holomorphic at* ∞.

We write $\mathcal{M}_k(\Gamma)$ *the set of all modular forms of weight* k *which forms a* $\mathbb{C}-$*vector space. We denote by*

$$\mathcal{M}(\Gamma) = \bigoplus_{k\in\mathbb{Z}} \mathcal{M}_k(\Gamma)$$

the set of all modular forms of any weight with respect to Γ. *It forms a graded ring.*

The last condition must be explained in some detail. There exists a period $h \in \mathbb{Z}$ such that $f(z+h) = f(z)$ (this follows from condition 2) which means we can write a Fourier expansion at ∞

$$f(z) = \sum_{n\in\mathbb{Z}} a_n q_h^n, \quad q_h = e^{\frac{2\pi i z}{h}}.$$

We say f is holomorphic at ∞ if $a_n = 0$ for $n < 0$.

Definition 1.1.6. *If* f *is a modular form of weight* $k \in \mathbb{Z}$ *and*

$$f(z) = \sum_{n\in\mathbb{N}} a_n q^n$$

is its Fourier expansion at ∞ *we say* f *is a* ***cusp form of weight*** k *if* $a_0 = 0$. *We write* $\mathcal{S}_k(\Gamma)$ *the set of all cusps forms of weight* k *which forms a vector subspace of* $\mathcal{M}_k(\Gamma)$. *We denote by*

$$\mathcal{S}(\Gamma) = \bigoplus_{k\in\mathbb{Z}} \mathcal{S}_k(\Gamma)$$

the set of cusps forms of any weight with respect to Γ. *It is a maximal ideal of* $\mathcal{M}(\Gamma)$.

Remark 1.1.7. $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) = \{0\}$ *for any odd* k.

Proposition 1.1.8 (Cf. [DS05] Theorem 3.5.1. and Theorem 3.6.1.)**.** *For any* k *and any congruence subgroup* Γ, $\mathcal{M}_k(\Gamma)$ *is a finite dimensional vector space.*

1.2 Examples

Example 1.2.1. *Take* $k \in \mathbb{Z}$ $k \geq 2$. *The* ***Eisenstein series of weight k*** *is defined by*

$$G_k(z) = {\sum_{(c,d)}}' \frac{1}{(cz+d)^k}, \quad z \in \mathbb{H},$$

where the primed summation means we sum over all integral pairs (c,d) *but* $(0,0)$. $G_k(z)$ *is a modular form of weight* k *for the full modular group. Note that it equals 0 if* $k > 2$ *is odd.*

Definition 1.2.2. *Let k and n be positive integers. We define*

$$\sigma_k(n) = \sum_{d|n} d^k.$$

Proposition 1.2.3 (Cf. [DS05] 1.1)**.** *For any $k > 2$*

$$G_k(z) = 2\zeta(k) - 2\frac{(2\pi i)^k}{(k-1)!}\sum_{n\geq 1}\sigma_{k-1}(n)q^n, \quad q = e^{2\pi iz}.$$

Where $\zeta(s)$ is the Riemann zeta function.

Definition 1.2.4. *We will denote by $E_k(z)$ the normalised Eisenstein series, i.e.*

$$E_k(z) = \frac{G_k(z)}{2\zeta(k)}.$$

This series has its first coefficient equal to 1.

Lemma 1.2.5 (Cf.[DS05] 1.2)**.** *The Eisenstein series of weight 2 G_2 satisfies*

$$G_2(\gamma(z)) = (cz+d)^2 G_2(z) - 2\pi ic(cz+d), \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

And then, it is not modular.

Example 1.2.6. *The **discriminant function** is defined by*

$$\Delta(z) = q\prod_{r=1}^{\infty}(1-q^r)^{24}, \quad q = e^{2\pi iz}$$

for all $z \in \mathbb{H}$. It is a cusp form of weight 12 for the full modular group.

Proof.

$$\begin{aligned}\frac{\Delta'(z)}{\Delta(z)} =& \frac{d}{dz}(\log(\Delta(z))) = \frac{d}{dz}(\log(q\prod_{r=1}^{\infty}(1-q^r)^{24})) = \frac{d}{dz}(\log(q) + \sum_{r=1}^{\infty}\log((1-q^r)^{24}) \\ =& \frac{d}{dz}\left(2\pi iz + 24\sum_{r=1}^{\infty}\log(1-q^r)\right) = 2\pi i - 48\pi i\sum_{r=1}^{\infty}\frac{rq^r}{1-q^r} = -48\pi i\left(\frac{-1}{24} + \sum_{n=1}^{\infty}\sigma_1(n)q^n\right) \\ =& \frac{6i}{\pi}G_2(z).\end{aligned}$$

From 1.2.5 we deduce

$$\frac{1}{(cz+d)^2}\frac{\Delta'(\gamma(z))}{\Delta(\gamma(z))} = \frac{\Delta'(z)}{\Delta(z)} + 12\frac{c}{cz+d},$$

for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, which is equivalent to

$$\frac{d}{dz}\left(\log(\Delta(\gamma(z))\right) = \frac{d}{dz}\left(\log(\Delta(z)(cz+d)^{12})\right),$$

since $\gamma'(z) = \frac{1}{(cz+d)^2}$. Integrating in both sides we have

$$\Delta(\gamma(z)) = w\Delta(z)(cz+d)^{12}, \quad w \in \mathbb{C}.$$

Taking some appropriate γ it is easy to see that, in fact, $w = 1$ and that Δ is a cusp form of weight 12.

If we develop the product we get the Fourier series of Δ (cf. [Leh43])

$$\Delta(z) = q - 24q^2 + 252q^3 - 1472q^4 + 4830q^5 - 6048q^6 - 16744q^7 + 84480q^8 - 113643q^9 - 115920q^{10} + \dots.$$

□

Definition 1.2.7. *We call the* ***Ramanujan function*** *to*

$$\tau : \mathbb{N} \longrightarrow \mathbb{Z}$$

defined by $\tau(n) = a_n$, *where* $\Delta(z) = \sum_{n\geq 1} a_n q^n$.

Remark 1.2.8. *We will see that this function is indeed multiplicative.*

Proposition 1.2.9. *Taking* $g_2(z) = 60G_4(z)$ *and* $g_3(z) = 140G_6(z)$,

$$\Delta(z) = (g_2(z))^3 - 27(g_3(z))^2.$$

Definition 1.2.10. *We call the Bernoulli numbers* B_k *the numbers such that*

$$\frac{t}{1-e^t} = \sum_{k\geq 0} B_k \frac{t^k}{k!}.$$

Lemma 1.2.11 (Cf. [DS05] Exercise 1.1.7)**.** *If* $\zeta(s)$ *is the Riemann zeta function, then*

$$2\zeta(k) = -\frac{(2\pi i)^k}{k!} B_k,$$

for all $k \geq 2$ *even.*

Remark 1.2.12. *From now on, we will rescale the Eisenstein series because it will be more convenient. We will divide by*

$$-2\frac{(2\pi i)^k}{(k-1)!}.$$

It is not difficult now to see that

$$G_k = -\frac{B_k}{2k} + \sum_{n\geq 1} \sigma_{k-1}(n) q^n,$$

where B_k *is the* k*-th Bernoulli number.*

1.3 Decomposition of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$

The goal of this section is to prove a special case of a more general fact about the space of modular forms which is, that for any $k > 2$, we have

$$\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) = \langle G_k \rangle \oplus \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z})).$$

Proof. Take f, g two different modular forms of weight $k \in \mathbb{Z}$ which are not cusps forms. Then, they have a q-expansion

$$f = \sum_{n=0}^{\infty} a_n q^n, \quad g = \sum_{n=0}^{\infty} b_n q^n.$$

Since they are not cusps $a_0 b_0 \neq 0$. Define $\lambda_0 = \frac{a_0}{b_0}$ and consider

$$f - \lambda_0 g = \sum_{n=0}^{\infty} (a_n - \lambda_0 b_n) q^n = \sum_{n=1}^{\infty} (a_n - \lambda_0 b_n) q^n,$$

because $a_0 - \lambda_0 b_0 = 0$. Hence, $f - \lambda_0 g \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$. Take $\{s_1, \ldots, s_m\}$ a basis of $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ and $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ such that

$$f - \lambda_0 g = \lambda_1 s_1 + \cdots + \lambda_m s_m,$$

equivalently

$$f = \lambda_0 g + \lambda_1 s_1 + \cdots + \lambda_m s_m.$$

In particular, $\{g, s_1, \ldots, s_m\}$ generate $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ for any g which is not a cusp form. Take now complex numbers $\mu_0, \ldots, \mu_m$ such that

$$\mu_0 g + \mu_1 s_1 + \cdots + \mu_m s_m = 0,$$

taking the Fourier series of this sum, gives us the identity $\mu_0 b_0 = 0$ (because $s_1, \ldots, s_m$ are cusps), and we had established $a_0 \neq 0$, so $\mu_0 = 0$. And now we have a trivial linear combination of the elements of the basis of $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$, so $\mu_1 = \cdots = \mu_m = 0$. This proves

$$\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) = \langle g \rangle \oplus \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z})),$$

as g is any modular form which is not a cusp, we can take $g = G_k$. □

1.4 Modular curves

In this section we will explain how modular forms with respect to congruence subgroups give rise to objects that can be made into Riemann surfaces passing through a process of compactification.

Definition 1.4.1. *Let $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ a congruence subgroup, we define the* ***modular curve*** *$Y(\Gamma)$ by*

$$Y(\Gamma) = \Gamma\backslash\mathbb{H} = \{\Gamma z : \ z \in \mathbb{H}\}.$$

In order to define a topology in $Y(\Gamma)$ take the subspace topology for $\mathbb{H}$ inherited from the Euclidean in $\mathbb{C}$. Then, the natural surjection

$$\begin{array}{rccc} \pi : & \mathbb{H} & \longrightarrow & Y(\Gamma) \\ & z & \longmapsto & \pi(z) = \Gamma z, \end{array}$$

which gives to $Y(\Gamma)$ the quotient topology.

Proposition 1.4.2. *For any congruence subgroup $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$, $Y(\Gamma)$ is Hausdorff.*

Proof. It is a result of Lie Group theory that proving that the quotient of a space by a Lie group action is Hausdorff it is only needed that the action is proper, but Γ is discrete, so properness is the same that proper discontinuity, i.e., $z, w \in \mathbb{H}$ there are U, V neighbourhoods of z and w respectively such that for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, if $\gamma(U) \cap V \neq \varnothing$, then $\gamma(z) = w$.

For a proof of this fact cf. [DS05] Proposition 2.1.1. □

Proposition 1.4.3. *$Y(\Gamma)$ has structure of differential manifold.*

Proof. By Lie Group theory, if the action was free (i.e., points were fixed only by the identity) we could say this with no more comment. The fact is that modular curves have points whose isotropy groups (i.e. subgroups of Γ fixing the point) are not always trivial. These points are called **elliptic points**, but their isotropy groups are finite and cyclic which even if it is a problematic situation we can endow holomorphic charts and give a differential structure.

The complete proof can be found in [DS05] Section 2.2. □

Theorem 1.4.4. *For any congruence subgroup $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$, $Y(\Gamma)$ can be compactified into a Riemannian curve $X(\Gamma)$. This curve can be described as the locus of points of some polynomials whose coefficients are rational.*

Proof. The compactification passes through extending the action of Γ from $\mathbb{H}$ to $\mathbb{H}^* = \mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$, in the natural way. This induces classes on $\mathbb{Q} \cup \{\infty\}$ which is invariant under the action, and these classes are called **cusps**. $X(\Gamma)$ is Hausdorff, compact, connected and has a structure of holomorphic Riemann surface. We don't define charts on it, as we didn't do for $Y(\Gamma)$ because we will not study further these curves. The complete proof can be found in [DS05] section 2.4.

Moreover, any compact Riemann surface is defined by the locus of points of a polynomial with coefficients in $\mathbb{C}$, by Modularity Theorem we know that there is a polynomial with coefficients in $\mathbb{Q}$ defining this curve.

□

This last theorem is important because the genus g of $X(\Gamma)$ gives information about the dimension of $\mathcal{S}_k(\Gamma)$ and gives information about congruences between modular forms.

Definition 1.4.5. *For $\Gamma(N), \Gamma_1(N), \Gamma_0(N)$ we write*

$$Y(N), Y_1(N), Y_0(N), \quad \text{and } X(N), X_1(N), X_0(N) \quad \text{respectively.}$$

1.5 Dirichlet characters

We state here multiple results about Dirichlet characters, these are basics on Number Theory and we do not give proofs of them. However, they can be found in many introductory texts in Number Theory (see for instance [Coh07]).

Definition 1.5.1. *Let N be a positive integer, a **Dirichlet character modulo** N is a homomorphism*

$$\chi : (\mathbb{Z}/N\mathbb{Z})^* \longrightarrow \mathbb{C}^*.$$

The set of all Dirichlet characters modulo N is denoted by $\mathrm{Hom}((\mathbb{Z}/N\mathbb{Z})^*, \mathbb{C}^*)$ and forms a group with the multiplication. It is called the **dual** of $(\mathbb{Z}/N\mathbb{Z})^*$. Take any χ a Dirichlet character and any non trivial element of $a \in (\mathbb{Z}/N\mathbb{Z})^*$. Then, by Euler theorem if $N' = \varphi(N)$ where φ is the Euler's totient function, then

$$1 = \chi(1) = \chi(a^{N'}) = (\chi(a))^{N'},$$

so the image of $(\mathbb{Z}/N\mathbb{Z})^*$ by χ is a subgroup of the N'-th roots of unity for any χ.

In order to understand these characters we can give some basic results in a more general context which we will not prove.

Proposition 1.5.2 (Cf. [Coh07] Proposition 2.1.16.)**.** *Let G be an abelian group, then $\mathrm{Hom}(G, \mathbb{C}^*) \simeq G$ and $G \simeq \mathrm{Hom}(\mathrm{Hom}(G, \mathbb{C}^*), \mathbb{C}^*)$ canonically.*

Proposition 1.5.3 (Cf.[Coh07] Corollary 2.1.33.)**.** *Let G be an abelian finite group of order n,*

1. *if $\chi \in \mathrm{Hom}(G, \mathbb{C}^*)$, then*

$$\sum_{g \in G} \chi(g) = \begin{cases} n & \text{if } \chi = \mathbb{1}, \\ 0 & \text{otherwise.} \end{cases}$$

2. *if $g \in G$, then*

$$\sum_{\chi \in \mathrm{Hom}(G, \mathbb{C}^*)} \chi(g) = \begin{cases} n & \text{if } g = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Remark 1.5.4. *We will write G_n instead of $(\mathbb{Z}/n\mathbb{Z})^*$ and $\widehat{G}_n$ instead of $\mathrm{Hom}(G_n, \mathbb{C}^*)$.*

Remark 1.5.5. *Let n, m be integers such that $n \mid m$, then there exists a epimorphism called reduction r from G_m to G_n, then if $\chi \in \widehat{G}_n$, then*

$$G_m \xrightarrow{r} G_n \xrightarrow{\chi} \mathbb{C}^*,$$

so $\chi \circ r \in \widehat{G}_m$.

Proposition 1.5.6. *Let n be an integer and d_1, d_2 two divisors. Let $\chi \in \widehat{G}_n$, if r_1, r_2 are the respectively reduction morphisms and $\chi_1 \in \widehat{G}_{d_1}, \chi_2 \in \widehat{G}_{d_2}$ are such that the diagram*

$$\begin{array}{ccc} G_n & \xrightarrow{r_1} & G_{d_1} \\ {\scriptstyle r_2}\downarrow & \overset{\chi}{\searrow} & \downarrow{\scriptstyle \chi_1} \\ G_{d_2} & \xrightarrow[\chi_2]{} & \mathbb{C}^* \end{array}$$

commutes, then there exists $\chi' \in \widehat{G}_d$ where $d = \gcd(d_1, d_2)$ such that if $r : G_n \longrightarrow G_d$ is the reduction morphism $\chi = \chi' \circ r$.

This is an easy result.

Corollary 1.5.7. *For any $\chi \in \widehat{G}_n$ there exists d dividing n which is the lowest integer such that χ is the composition of a Dirichlet character modulo d and the respective reduction morphism. We call this integer the **conductor** of χ and say that $\chi \in \widehat{G}_n$ is primitive if it has conductor n.*

Now we are able to define in a proper way the modular forms with non-trivial nebentype.

Definition 1.5.8. *Let N, k be integers, let χ be a Dirichlet character modulo N and $f \in \mathcal{M}_k(\Gamma_1(N))$. We say that f has **nebentype** χ if,*

$$f|[\gamma]_k(z) = \chi(d) f(z), \qquad \text{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N),\ z \in \mathbb{H}.$$

The set of all such modular forms is denoted by

$$\mathcal{M}_k(N, \chi).$$

Proposition 1.5.9 (Cf. [DS05] Exercise 4.3.4)**.** *Let N, k be integers, then*

$$\mathcal{M}_k(\Gamma_1(N)) = \bigoplus_{\chi \in \widehat{G}_N} \mathcal{M}_k(N, \chi).$$

Remark 1.5.10. *Note that if we take $\chi = \mathbb{1}$, then*

$$\mathcal{M}_k(N, \chi) = \mathcal{M}_k(N, \mathbb{1}) = \mathcal{M}_k(\Gamma_0(N)).$$

1.6 Modular forms mod ℓ

Definition 1.6.1. *Let f be a modular form for the full modular group. We define the operator*

$$\theta = q\frac{d}{dq},$$

i.e.,

$$\theta(\sum_{n\geq 0} a_n q^n) = \sum_{n\geq 1} n a_n q^n.$$

Lemma 1.6.2 (Cf. [SD73] §3, Lemma 3.)**.** *If $f \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ then, $(12\theta f - kE_2 f) \in \mathcal{M}_{k+2}(\mathrm{SL}_2(\mathbb{Z}))$ and $(12\theta E_2 - E_2^2) \in \mathcal{M}_4(\mathrm{SL}_2(\mathbb{Z}))$.*

The proof is pure calculation.

Using the fact that any modular form for the full modular group can be written as an isobaric polynomial in E_4 and E_6 we can reformulate this lemma in terms of a derivation in a graded algebra.

Lemma 1.6.3 (Cf. [SD73] §3 Corollary of Lemma 3)**.** *$\partial = 12\theta - kE_2$ is determined as the derivation in the graded algebra of modular forms such that $\partial E_4 = -4E_6$ and $\partial E_6 = -6E_4^2$.*

Definition 1.6.4. *For any ℓ prime, $S = \mathbb{Z} \setminus (\ell)$ is a multiplicative system and we can define the **local ring at** ℓ by $\mathfrak{o} = S^{-1}\mathbb{Z}$. Now, we define $\mathfrak{M}_k$ the $\mathfrak{o}$-module of modular forms of weight k whose Fourier coefficients are all in $\mathfrak{o}$. Then,*

$$\overline{\mathfrak{M}}_k = \left\{ \overline{f} = \sum_{n\geq 0} \overline{a_n} q^n : \ f = \sum_{n\geq 0} a_n q^n \in \mathfrak{M}_k \right\},$$

where the line above means reduction modulo ℓ.

Define

$$\overline{\mathfrak{M}} = \sum_{k\in\mathbb{Z}} \overline{\mathfrak{M}}_k.$$

Remark 1.6.5. *Note that we cannot write the direct sum symbol because it would mean that there are no congruences.*

Definition 1.6.6. *Let $\overline{f} \in \overline{\mathfrak{M}}$, we define the filtration*

$$\omega(\overline{f}) = \inf\left\{k : \ \overline{f} \in \mathfrak{M}_k\right\}.$$

Lemma 1.6.7 (von Staudt's Theorem)**.**

1. *If $2k \equiv 0 \pmod{\ell - 1}$ then, $\ell B_{2k} \equiv -1 \pmod{\ell}$.*
2. *If $2k \not\equiv 0 \pmod{\ell - 1}$ then, $v_\ell\left(\frac{B_{2k}}{2k}\right) \geq 0$ and the residue class modulo ℓ only depends on the residue class of $2k$ modulo $\ell - 1$.*

Proof. Cf. [BC67] *Théorème 4* pp. 431-433. □

Lemma 1.6.8 (Ramanujan)**.** *Let $f \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ then there is a unique $\phi \in \mathbb{Q}[X,Y]$ a polynomial such that $\phi(E_4, E_6) = f$ and ϕ is isobaric in E_4 and E_6.*

Let ℓ be a prime number and $\mathfrak{o}$ its local ring. If $f = \sum_{n\geq 0} a_n q^n$ for $a_n \in \mathfrak{o}$, then $\phi \in \mathfrak{o}[X,Y]$.

Proposition 1.6.9. *Let $\ell > 3$ be a prime number and $A, B \in \mathfrak{o}[X,Y]$ such that*

$$A(E_4, E_6) = E_{\ell-1}, \quad B(E_4, E_6) = E_{\ell+1}.$$

Then, if we denote by an $\overline{A}, \overline{B}$ their class modulo ℓ we have

1. *$\overline{A}(\overline{E_4}, \overline{E_6}) = 1$, and $\overline{B}(\overline{E_4}, \overline{E_6}) = \overline{E_2}$.*
2. *$\partial\overline{A}(\overline{E_4}, \overline{E_6}) = \overline{B}(\overline{E_4}, \overline{E_6})$ and $\partial\overline{B}(\overline{E_4}, \overline{E_6}) = -\overline{E_4}\overline{A}(\overline{E_4}, \overline{E_6})$.*
3. *$\overline{A}(\overline{E_4}, \overline{E_6})overline{E_6})$ has no repeated factors and $(\overline{A}(\overline{E_4}, \overline{E_6}), \overline{B}(\overline{E_4}, \overline{E_6})) = (1)$.*
4. *$\overline{\mathfrak{M}}$ is naturally isomorphic to $\mathbb{F}_\ell/(\overline{A} - 1)$ and has a natural grading structure with values in $\mathbb{Z}/(\ell-1)\mathbb{Z}$.*

Proof. 1. By the above lemma $\ell B_{\ell-1} \equiv -1 \pmod{\ell}$ which means $0 = v_\ell(\ell B_{\ell-1}) = v_\ell(B_{\ell-1}) + 1$ so $v_\ell(B_{\ell-1}) = -1$, in particular, since $\ell \nmid \ell - 1$ we have that

$$E_{\ell-1} = 1 + \frac{2\ell-2}{B_{\ell-1}} \sum_{n\geq 0} \sigma_{\ell-2}(n) q^n$$

so A has coefficients in $\mathfrak{o}$ and $A \equiv 1 \pmod{\ell}$ since $v_\ell(\frac{2\ell-2}{B_{\ell-1}}) = 1$.

From the second part of the lemma $B_{\ell+1}/(\ell+1) \equiv B_2/2 \equiv -1/12 \pmod{\ell}$, so $E_{\ell+1} \in \mathfrak{M}_{\ell+1}$, and $B \in \mathfrak{o}[X,Y]$.

Now, by the little Fermat's theorem $d^\ell \equiv d \pmod{\ell}$ for any integer d, so $\sigma_\ell(n) = \sum_{d|n} d^\ell \equiv \sum_{d|n} d = \sigma_1(n)$, and the second part follows. The first part follows from the above lemma.

2. From point 1, we know that $\theta(\overline{A}(\overline{E_4}, \overline{E_6})) = 0$ so

$$\partial(\overline{A}(\overline{E_4}, \overline{E_6})) = -(\ell-1)\overline{E_2}\overline{A}(\overline{E_4}, \overline{E_6}) = \overline{E_2} = \overline{B}(\overline{E_4}, \overline{E_6}).$$

So $\partial A - B$ has a q-expansion where all its coefficients are divisible by ℓ and since it is a modular form of weight $\ell + 1$ it is in $\ell\mathfrak{o}[\overline{E}_4, \overline{E}_6]$. Hence $\partial\overline{A} = \overline{B}$.

If we follow the same process for B we get

$$\partial\overline{B}(\overline{E}_4, \overline{E}_6) = (12\theta - \overline{E}_2)\overline{B}(\overline{E}_4, \overline{E}_6) = (12\theta - \overline{E}_2)\overline{E}_2$$

which is a modular form of weight 4 by lemma 1.6.2. Calculations show that it equals $-\overline{E}_4 = -\overline{E}_4\overline{A}(\overline{E}_4, \overline{E}_6)$. The argument to show that $\partial\overline{B} = -\overline{E}_4\overline{A}$ is the same as before and we do not repeat it.

3. Suppose now that $\overline{A}(X,Y)$ is exactly divisible by some factor of the form $(X - cY)^n$ with $n > 0$ and $c \neq 0$ in the algebraic closure of $\mathbb{F}_\ell$. We know that $E_4^3 - E_6^2$ has a zero constant term, so does its reduction modulo ℓ and $\overline{A}(\overline{E}_4, \overline{E}_6)$ has not (because if it had, it would be divisible either by $\overline{E}_4$ or $\overline{E}_6$ which are not invertible, but $\overline{A}(\overline{E}_4, \overline{E}_6) = 1$), then $c \neq 0$. Take $\partial(X - cY) = 12(c-1)X^2Y$, so since $\ell \nmid 12$ and $c \neq 1$ we have that $X^3 - cY^2$ has double factors.

   Now, since $\partial\overline{A} = \overline{B}$, $\overline{B}$ is divisible by $\overline{E}_4^3 - c\overline{E}_6^2$ exactly $n-1$ times, and hence, if $n > 1$ $\partial\overline{B} = -\overline{E}_4\overline{A}$ and since $\overline{E}_4$ is not divisible by $\overline{E}_4^3 - c\overline{E}_6^2$ we deduce that $\overline{A}$ is divisible by $\overline{E}_4^3 - c\overline{E}_6^2$ exactly $n-2$ times, against our hypothesis.

   Assume $\overline{E}_4$ divides $\overline{A}$ exactly n times. Then $\overline{E}_4$ divides exactly $n-1$ times $\overline{B}$ by the same argument than the case before. Now $\partial\overline{B} = \overline{E}_4\overline{A}$, which means that $\partial\overline{B}$ is divisible by $\overline{E}_4$ exactly $n-2$ times and at the same time $n+1$ times, which is contradictory. And, $\partial\overline{E}_4 = -4\overline{E}_6$ which is coprime to $\overline{E}_4$ since $2 \nmid \ell$.

   The argument for powers of $\overline{E}_6$ is almost the same and we do not make it because it would be repetitive.

   Summarising we proved that $\overline{A}$ has no repeated factors and their factors do not divide $\overline{B}$ (the cases above are the only possible since they are the only isobaric polynomials in $\overline{E}_4$ and $\overline{E}_6$).

4. Let $\mathfrak{a}$ be the kernel of the map

$$\varphi : \mathbb{F}_\ell[X,Y] \longrightarrow \mathbb{F}_\ell[[q]]$$

   which sends isobaric polynomials in $\overline{E}_4, \overline{E}_6$ to their reduced q-expansion. The inclusion $(\overline{A}-1) \subseteq \ker\varphi$ is clear. $\mathbb{F}_\ell[\overline{E}_4, \overline{E}_6]$ has dimension 2, so in order to prove the other inclusion we only need to see that $\overline{A}-1$ is irreducible as a polynomial in X, Y. Let

$$F(X,Y) = F_n(X,Y) + F_{n-1}(X,Y) + \cdots + 1,$$

   be an irreducible factor, where $F_i(X,Y)$ is isobaric of weight i. Let ζ be a primitive $(\ell-1)$-th root of unity in $\mathbb{F}_\ell$, then $F(\zeta^2 X, \zeta^3 Y)$ is also a factor of $\overline{A}-1$. So, considering the terms of highest weight we see that $F_n(X,Y)^2$ divides $\overline{A}$, but this is a contradiction, since we have seen in point 3 that $\overline{A}$ has no repeated factors.

□

For further results, we will need a simple but powerful tool which is this lemma, where we state the behaviour of the filtration defined for $\overline{\mathfrak{M}}$.

Lemma 1.6.10 (Katz)**.** *Let ℓ be a prime number and let f be a modular form of weight k, and let $F \in \mathfrak{o}[X,Y]$ such that $f = F(E_4, E_6)$. Assume $\overline{f} \neq 0$. Then $\omega(\overline{f}) < k$ if, and only if, $\overline{A}$ divides $\overline{F}$.*

Let $\overline{f} \in \overline{\mathfrak{M}}$, then $\omega(\theta\overline{f}) \leq \omega(\overline{f}) + \ell + 1$, with equality if, and only if, $\omega(\overline{f}) \not\equiv 0 \pmod{\ell}$.

Proof. In order to prove the first part suppose that $\overline{F} = \overline{R_1} \dots \overline{R_s}$ for some irreducible factors. Since $\omega(\overline{f}) < k$ and k is the weight of f some factor $\overline{R_i}(\overline{E}_4, \overline{E}_6)$ must be equal to 1, so $\overline{R_i} = 1 = \overline{A}$. Conversely, if $\overline{A}$ divides $\overline{F}$ then it is clear that the modular form $\overline{f}$ decreases. This result uses the above theorem.

Let k be the graduation of $\overline{f}$ in $\overline{\mathfrak{M}}$ and let $f = F(E_4, E_6)$ a modular form of weight k whose reduction is $\overline{f}$. Using

$$12\theta\overline{f} = \overline{A}(\overline{E}_4, \overline{E}_6)\delta\partial\overline{F}(\overline{E}_4, \overline{E}_6) + k\overline{B}(\overline{E}_4, \overline{E}_6)\overline{f},$$

which means, $12\theta\overline{f}$ is the image of $\overline{A}\partial\overline{F} + k\overline{BF}$. And, by the first part, $\overline{A} \nmid \overline{F}$ because $\overline{f} \neq 0$, so by the above theorem $\overline{A}\partial\overline{F} + k\overline{BF}$ divides $\overline{A}$ if $k \equiv 0 \pmod{\ell}$ and the result follows.

□

The proof of this proposition is not difficult and we do not reproduce it.

Proposition 1.6.11. *If $\ell = 2, 3$ then $\overline{E}_4 = \overline{E}_6 = \overline{E}_2 = 1$ and $\overline{\mathfrak{M}} = \mathbb{F}_\ell[\overline{\Delta}]$. There is no grading and $\partial\overline{\mathfrak{M}} = 0$.*

Chapter 2

Hecke Operators

Hecke operators are a very strong tool we have that allows us to study deeply the structure of the space of modular forms. In particular, it will be very useful when we try to find where the Fourier coefficients lie, and if they are somehow algebraic, the fact is that the space of modular forms can be spanned by modular forms with algebraic coefficients. This is very important because it justifies why we do Number Theory with these forms and why we try to find congruences between their Fourier coefficients.

2.1 The double coset operator

Definition 2.1.1. *Let G be a group, $H, K \subseteq G$ subgroups and $g \in G$. A* ***double coset in*** *G* ***of*** *g is the set*

$$HgK.$$

Proposition 2.1.2 (Cf. [DS05] Lemma 5.1.1. and Lemma 5.1.2.)**.** *Let Γ_1, Γ_2 be two congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ and $\alpha \in \mathrm{GL}_2(\mathbb{Q})^+$. The orbit space $\Gamma_1 \backslash \Gamma_1 \alpha \Gamma_2$ is a finite disjoint union of orbits.*

Definition 2.1.3. *Let Γ_1, Γ_2 be two congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ and $\alpha \in \mathrm{GL}_2(\mathbb{Q})^+$. The $\Gamma_1 \alpha \Gamma_2$* ***operator of weight*** *k is defined by*

$$f[\Gamma_1 \alpha \Gamma_2]_k = \sum_i f|[\alpha_i]_k,$$

where $\{\alpha_i\}$ are orbit representatives, i.e., $\Gamma_1 \alpha \Gamma_2 = \bigcup_i \Gamma_1 \alpha_i$ and $\Gamma_1 \alpha_i \cap \Gamma_1 \alpha_j = \varnothing$ if $i \neq j$.

Proposition 2.1.4 (Cf. [Miy06] 2.7)**.** *The above definition is independent from the choice of the orbit representatives and takes modular forms with respect to Γ_1 to modular forms with respect to Γ_2. It also takes cusp forms with respect to Γ_1 to cusp forms with respect to Γ_2.*

What we get is an operator which transforms a modular form with respect to Γ_1 to a modular form with respect to Γ_2 and respecting the subspace of cusps.

$$[\Gamma_1 \alpha \Gamma_2]_k : \mathcal{M}_k(\Gamma_1) \longrightarrow \mathcal{M}_k(\Gamma_2).$$

Proposition 2.1.5 ([DS05] p. 166.)**.** *Let Γ_1, Γ_2 be as above. Then,*

(i) If $\Gamma_2 \subseteq \Gamma_1$ and we take $\alpha = Id$, then $f[\Gamma_1 \alpha \Gamma_2]_k = f$. The operator induces the natural inclusion from $\mathcal{M}_k(\Gamma_1)$ to $\mathcal{M}_k(\Gamma_2)$.

(ii) If $\alpha^{-1}\Gamma_1\alpha = \Gamma_2$ then $f[\Gamma_1 \alpha \Gamma_2]_k = f|[\alpha]_k$, which induces an isomorphism from $\mathcal{M}_k(\Gamma_1)$ to $\mathcal{M}_k(\Gamma_2)$.

(iii) If $\Gamma_1 \subseteq \Gamma_2$ and we take $\alpha = Id$, then $[\Gamma_1 \alpha \Gamma_2]_k$ is the projection of $\mathcal{M}_k(\Gamma_1)$ onto its subspace $\mathcal{M}_k(\Gamma_2)$, which is a surjection.

2.2 Diamond operator

In the first chapter we have mentioned that for any positive integer N we have

$$\Gamma_0(N)/\Gamma_1(N) \simeq (\mathbb{Z}/N\mathbb{Z})^*,$$

(cf. Proposition 1.1.3).

In order to introduce the Diamond operator we will use that result.

Recall that we defined

$$\begin{array}{ccc} \Gamma_0(N) & \longrightarrow & (\mathbb{Z}/N\mathbb{Z})^* \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \longmapsto & d \end{array}$$

This was obviously a surjection since the determinant of the matrices in $\gamma \in \Gamma_0(N)$ are such that $\det(\gamma) \equiv ad \pmod N$ and $\det(\gamma) \equiv 1 \pmod N$.

Its kernel is formed by all matrices with $d = 1$ and, because of $ad \equiv 1 \pmod N$ we have that

$$\ker = \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) : \ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod N \right\} = \Gamma_1(N).$$

Taking a double coset operator with $\alpha \in \Gamma_0(N)$ and $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$, we have the second case in Proposition 2.1.5, so

$$f[\Gamma_1(N)\alpha\Gamma_1(N)] = f|[\alpha]_k.$$

Definition 2.2.1. *With the above notation, we define the* ***Diamond operator*** *to be the double coset operator with* $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$ *and* $\alpha \in \Gamma_0(N)$*. We denote it by*

$$\begin{array}{rccc} \langle d \rangle : & \mathcal{M}_k(\Gamma_1(N)) & \longrightarrow & \mathcal{M}_k(\Gamma_1(N)) \\ & f & \longmapsto & \langle d \rangle f = f|[\alpha]_k \end{array}$$

for any $\alpha = \begin{pmatrix} a & b \\ c & \delta \end{pmatrix}$ *with* $\delta \equiv d \pmod{N}$.

2.3 T_p and T_n operators

Definition 2.3.1. *Take* $p \in \mathbb{Z}$ *a prime number, and* $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$*. Take* N *a positive integer. We define the* T_p *operator to be*

$$T_p : \mathcal{M}_k(\Gamma_1(N)) \longrightarrow \mathcal{M}_k(\Gamma_1(N))$$

defined by $T_p f = [\Gamma_1(N)\alpha\Gamma_1(N)]_k$.

Proposition 2.3.2. *Let* $N \in \mathbb{Z}$ *positive and* p *a prime number, a set of representatives of*

$$T_p : \mathcal{M}_k(\Gamma_1(N)) \longrightarrow \mathcal{M}_k(\Gamma_1(N))$$

is

- $\left\{ \begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix} \right\}_{j=0}^{p-1}$ *if* $p \mid N$,
- $\left\{ \begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix} \right\}_{j=0}^{p-1} \cup \left\{ \begin{pmatrix} m & n \\ N & p \end{pmatrix} \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \right\}$ *if* $p \nmid N$ *and* $mp - nN = 1$.

Proof. The proof of this fact is not difficult and is left to the reader. □

Proposition 2.3.3. *Let* N *be a fixed natural number, and* p *and* q *different prime numbers,* d, e *numbers prime to* N.

1. $\langle d \rangle T_p = T_p \langle d \rangle$.
2. $\langle d \rangle \langle e \rangle = \langle e \rangle \langle d \rangle = \langle de \rangle$.
3. $T_p T_q = T_q T_p$.

Proof. This is a direct application of the above Proposition. □

Proposition 2.3.4. *Let p be a prime number. Let f be a modular form of weight k with respect to $\Gamma_1(N)$ for a fixed N. Let $\{a_n(f)\}_{n\geq 0}$ the coefficients of its Fourier expansion.*

1. *Let p be a prime number and let $\mathbb{1}_N : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{C}^*$ the trivial character modulo N (i.e. $\mathbb{1}_N(m) = 1$ if $m \nmid N$ and 0 otherwise). Then,*

$$a_n(T_p f) = a_{np}(f) + \mathbb{1}_N(p)p^{k-1}a_{n/p}(\langle p\rangle f)$$

are the Fourier coefficients of $T_p(f)$.

2. *Let now $\chi : (\mathbb{Z}/N\mathbb{Z})^* \longrightarrow \mathbb{C}^*$ be a Dirichlet character modulo N. If $f \in \mathcal{M}_k(N,\chi)$ then also $T_p f \in \mathcal{M}_k(N,\chi)$, and now its Fourier expansion is*

$$T_p f(z) = \sum_{m=0}^{\infty} (a_{mp}(f) + \chi(p)p^{k-1}a_{m/p}(f))q^m.$$

Where we are considering $a_{n/p}(f) = 0$ if n/p is not a natural number.

Proof. For the first assertion, using 2.3.2 we calculate

$$f|\left[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}\right]_k (z) = p^{k-1}p^{-k} f\left(\frac{z+j}{p}\right) = p^{-1}\sum_{n\geq 0} a_n(f)e^{2\pi i n(z+j)/p} = p^{-1}\sum_{n\geq 0} a_n(f)q^{n/p}e^{2\pi ij/p}.$$

Hence,

$$\begin{aligned} T_p(f) &= \sum_{j=0}^{p-1} f|\left[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}\right]_k = \sum_{j=0}^{p-1} p^{-1}\sum_{n\geq 0} a_n(f)q^{n/p}e^{2\pi inj/p} = \sum_{n\geq 0} p^{-1}a_n(f)q^{n/p}\sum_{j=0}^{p-1} e^{2\pi ijn/p} \\ &= \sum_{p|n} a_n(f)q^{n/p} = \sum_{n\geq 0} a_{np}(f)q^n. \end{aligned}$$

Because both sums are absolutely convergent and $\sum_{j=0}^{p-1} e^{2\pi ijn/p} = p(1 - \mathbb{1}_n(p))$. This proves the case $p \mid N$. If $p \nmid N$ then we have to sum up with

$$\begin{aligned} f|\left[\begin{pmatrix} m & n \\ N & p \end{pmatrix}\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}\right]_k (z) &= (\langle p\rangle f)|\left[\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}\right]_k (z) = p^{k-1}(\langle p\rangle f)(pz) \\ &= p^{k-1}\sum_{n\geq 0} a_n(\langle p\rangle f)q^{np}, \end{aligned}$$

as we wanted to see. The second assertion follows from the first. □

T_p and $\langle d\rangle$ are a type of Hecke operator, but we want a more general definition which doesn't need p to be a prime number or d to be prime to N.

Definition 2.3.5. *Let N be a positive number. Take n any positive integer, if $\gcd(n,N) = 1$ $\langle n\rangle$ is properly defined in the second section. For $\gcd(n,N) > 1$ we define $\langle n\rangle = 0$.*

Remark 2.3.6. *It is obvious form this definition and the properties of the diamond operator that* $\langle nm\rangle = \langle n\rangle\langle m\rangle$.

Definition 2.3.7. *We want now to define* T_n *for arbitrary* n. *We will make it inductively,*

- *Take* $T_1 = 1$.
- *For prime* p T_p *is already defined.*
- *For* $n = p^r$, $r > 1$ *and* p *prime, put* $T_{p^r} = T_p T_{p^{r-1}} - p^{k-1}\langle p\rangle T_{p^{r-2}}$.
- *For* $n = p_1^{r_1} \dots p_s^{r_s}$ *take* $T_n = T_{p_1^{r_1}} \dots T_{p_s^{r_s}}$.

Remark 2.3.8. *Using the already seen properties of* T_p, *it is clear that* $T_{nm} = T_n T_m$, *for* $\gcd(n, m) = 1$.

Proposition 2.3.9. *Let* $f \in \mathcal{M}_k(\Gamma_1(N))$ *and* $\{a_m\}_m$ *its Fourier coefficients. Then, for any* n $\left\{\sum_{d|n,m} d^{k-1} a_{nm/d^2}\right\}_m$ *are the Fourier coefficients of* $T_n f$.

Proof. This is a direct consequence of Proposition 2.3.4 and Proposition 2.3.3. □

2.4 Eigenvalues of Hecke operators

Definition 2.4.1. *Let* N, k *be a positive integers and* f *a modular form of weight* k *with respect to* $\Gamma_1(N)$. *We say that* f *is a* ***Hecke form*** *if for all* n, *there exist* $\lambda_n \in C$ *such that*

$$T_n f = \lambda_n f.$$

Let $\{a_n\}_n$ *be its Fourier coefficients. We say that* f *is a* ***normalised Hecke form*** *if* $a_1 = 1$.

Proposition 2.4.2. *Let* f *be a normalised Hecke form of weight* k *with respect to* $\Gamma_1(N)$. *Let* $\{a_n\}_n$ *be its Fourier coefficients. Then, for all* n, m *coprime,*

$$a_n a_m = a_{nm}.$$

If $r \geq 1$, $a_{p^{r+1}} = a_p a_{p^r} - p^{k-1} a_{p^{r-1}}$.

Proof. We know that, for all n

$$T_n f = \sum_{m=0}^{\infty} \left(\sum_{d|n,m} d^{k-1} a_{nm/d^2} \right) q^m$$

and

$$T_n f = \lambda_n f = \sum_{m=0}^{\infty} a_m \lambda_n q^m.$$

Then,

$$\lambda_n a_m = \sum_{d|n,m} d^{k-1} a_{nm/d^2},$$

in particular

$$\lambda_n = \lambda_n a_1 = a_n, \quad \forall n.$$

Thus, if $\gcd(n,m) = 1$

$$a_n a_m = \lambda_n a_m = \sum_{d|1} d^{k-1} a_{nm/d^2} = a_{nm}.$$

This proves the first assertion, the second follows immediately putting $n = p$ and $m = p^r$ and

$$a_p a_{p^r} = \sum_{d|p,p^r} d^{k-1} a_{p^{r+1}/d^2} = \sum_{d=1,p} d^{k-1} a_{p^{r+1}/d^2} = a_{p^{r+1}} + p^{k-1} a_{p^{r-1}}.$$

□

Proposition 2.4.3. *Let $f \in \mathcal{M}_k(N,\chi)$. Then, f is a Hecke form if and only if $f = \sum_{n\geq 1} a_n q^n$ satisfy*

$$\begin{cases} a_1 = 1, \\ a_{p^r} = a_p a_{p^{r-1}} - \chi(p) p^{k-1} a_{p^{r_2}}, \text{ for all prime } p \text{ and } r \geq 2, \\ a_{mn} = a_n a_m, \text{ when } n, m \text{ are coprime.} \end{cases}$$

Proof. The Fourier coefficients of any Hecke form satisfies the three conditions by definition.

Assume now that f satisfy all conditions, then f is normalised and being an eigenform is equivalent to

$$a_m(T_p f) = a_p a_m,$$

for all prime p and $m > 0$. If p does not divide m, we know that

$$a_m(T_p f) = \sum_{d|(m,p)} \chi(d) d^{k-1} a_{mn/d^2} = a_{pm}$$

which equals $a_p a_m$ by the third condition.

Assume now $p|m$, then $m = p^r m'$ where $p \nmid m'$. By the above formula

$$a_m(T_p f) = a_{p^{r+1}m'} + \chi(p) p^{k-1} a_{p^{r-1}m'}$$

so by the third condition we get

$$a_m(T_p f) = \left(a_{p^{r+1}} + \chi(p) p^{k-1} a_{p^{r-1}}\right) a_{m'}$$

The second condition implies that

$$a_{p^{r+1}} + \chi(p)p^{k-1}a_{p^{r-1}} = a_p a_{p^r},$$

summarising

$$a_m(T_p f) = a_p a_{p^r} a_{m'} = a_p a_m.$$

□

2.5 Level N=1

In this section we want to prove some results for the level $N = 1$, i.e., $\Gamma_1(1) = \mathrm{SL}_2(\mathbb{Z})$. The results in this section will be generalised in next section for N arbitrary, but it is a nice exercise because it allows us to deepen our understanding in the theory of Hecke forms. The results in this section are slightly stronger for the case $N = 1$ than the results in the next section, we prove that Hecke eigenvalues are real numbers. The proof presented here could be rearranged to fit N arbitrary but we would face a problem determining the coefficients a_n for $(n, N) > 1$. In next section, given stronger results we overcome these difficulties.

Proposition 2.5.1. *If we take $\mathbb{H}\backslash\mathrm{SL}_2(\mathbb{Z})$, there exists a representative of each orbit in the set*

$$\mathcal{D} = \{z \in \mathbb{H} : \ |\Re(z)| \leq 1/2, |z| \geq 1\},$$

i.e., it is a fundamental domain of $\mathbb{H}$ under the action of $\mathrm{SL}_2(\mathbb{Z})$.

Theorem 2.5.2. *The Hecke forms in $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ form a basis in $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ for every k.*

In order to prove this theorem we need to define an inner product between modular forms.

Definition 2.5.3. *Define the **hyperbolic measure** in $\mathbb{H}$ as*

$$d\mu(x + iy) = \frac{dxdy}{y^2}.$$

Thus, for every $f, g \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$, we define

$$\langle f, g\rangle = \int_{\mathbb{H}\backslash\mathrm{SL}_2(\mathbb{Z})} f(z)\overline{g(z)}(\Im(z))^k d\mu(z),$$

*this product is called the **Petersson inner product.***

Remark 2.5.4. *This integral makes sense since $\Im(z)^k|f(z)|^2$ is a bounded function and the measure and the function are invariant under $\mathrm{SL}_2(\mathbb{Z})$. Also*

$$\int_{\mathbb{H}\backslash\mathrm{SL}_2(\mathbb{Z})} d\mu(z) = \frac{\pi}{3} < \infty.$$

Remark 2.5.5. *This product is linear in f, conjugate linear in g and positive definite.*

Remark 2.5.6. *The Petersson inner product can be generalised for all congruence subgroups. So, if Γ is a congruence subgroup and $f, g \in \mathcal{S}_k(\Gamma)$ then define*

$$\langle f, g\rangle = \int_{\mathbb{H}\backslash\Gamma} f(z)\overline{g(z)}(\Im(z))^k d\mu(z).$$

Proof. As we have seen in the first chapter, $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ splits as the direct sum of the space generated by cusps forms and the Eisenstein series G_k. So, first of all, let's see that G_k is a Hecke form. The only think we have to check is that the Fourier coefficients of G_k satisfy the identity

$$a_n a_m = \sum_{d|n,m} d^{k-1} a_{nm/d^2},$$

which is easy.

We will see now that the cusps forms have a basis of Hecke forms. For N, n such that $\gcd(n, N) = 1$ we have $\langle T_n f, g\rangle = \langle f, T_n g\rangle$ (see [Shi71] 3.4-3.5). As we treat $N = 1$ this condition is satisfied for any $n \in \mathbb{N}$. We also have seen that the T_n commute, and by linear algebra we know that $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ is spanned by all simultaneous eigenvector of all T_n. We have characterised all these forms by the condition on their Fourier coefficients

$$a_n a_m = \sum_{d|n,m} d^{k-1} a_{nm/d^2}.$$

Now,

$$a_n\langle f, f\rangle = \langle a_n f, f\rangle = \langle T_n f, f\rangle = \langle f, T_n f\rangle = \langle f, a_n f\rangle = \overline{a_n}\langle f, f\rangle,$$

so $\{a_n\}_n \subseteq \mathbb{R}$. If we take two Hecke forms in $f, g \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$, with Fourier coefficients $\{a_n\}_n$ and $\{b_n\}_n$ respectively, then

$$a_\langle f, g\rangle = \langle a_n f, g\rangle = \langle T_n f, g\rangle = \langle f, T_n g\rangle = \overline{b_n}\langle f, g\rangle = b_n\langle f, g\rangle.$$

As they are different, $a_n \neq b_n$ for at least one n, so $\langle f, g\rangle = 0$, and again by linear algebra, they are linearly independent, as we wanted to see.

□

Theorem 2.5.7. *The Fourier coefficients of a Hecke form $f \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ are real algebraic integers of degree less or equal to* $\dim(\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z})))$.

Proof. Consider a basis $f_1, \cdots, f_d$ of Hecke forms of $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ and consider the lattice generated by these elements, call it L. Then, $\mathrm{rank}_{\mathbb{Z}} L = d = \dim_{\mathbb{C}} \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$. We know that, for any n and for any $f(z) = \sum_{m\geq 0} a_m(f) q^m \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$

$$T_n f = \sum_{m\geq 0} \sum_{d|n,m} d^{k-1} a_{nm/d^2}(f).$$

Thus,

$$T_n L \subseteq L.$$

And we can express the action of T_n in L as a integral matrix of $d \times d$ entries. In particular, the eigenvalues of T_n are integral numbers of degree at most d because the characteristic polynomial of the matrix has degree d. These eigenvalues are exactly the Fourier coefficients of the basis $f_1, \ldots, f_d$ and we already know that they are real numbers, from which the result follows. □

2.6 Case N arbitrary

This section is philosophically important, since we want to deal with congruences between coefficients for Fourier series of modular forms. We could start talking about congruences and prove some results, but these would be meaningless if we don't see that we do have modular forms with integral coefficients.

In this section $G = \mathrm{Aut}_{\mathbb{Q}}(\mathbb{C})$.

Proposition 2.6.1 (Cf. [DI95] section 3.6.)**.** *$\mathcal{M}_k(N, \chi)$ is a vector space of finite dimension for any N and all χ. Hence, there exists a finite basis of Hecke forms.*

Lemma 2.6.2. *If $f \in \mathcal{M}_k(N, \chi)$ is an eigenform for T_m, then for all $\sigma \in G$, f^σ is also an eigenform in $\mathcal{M}_k(N, \chi^\sigma)$.*

Proof. Take $f = \sum_{m=0}^{\infty} a_m(f) q^m$, then $f^\sigma = \sum_{m=0}^{\infty} a_m^\sigma(f) q^m$. We only need to see that

$$a_m(f^\sigma) a_n(f^\sigma) = \sum_{d|(n,m)} \chi^\sigma(d) d^{k-1} a_{mn/d^2}(f^\sigma).$$

We know, for all m $a_m(f^\sigma) = a_m(f)^\sigma$, so since f is an eigenform for T_n then

$$a_n(f) a_m(f) = \sum_{d|(n,m)} \chi(d) d^{k-1} a_{nm/d^2}(f),$$

applying σ and using that σ respects product the result follows. □

Lemma 2.6.3. *For N and χ, if $f \in \mathcal{M}_k(N, \chi)$ then, $f^\sigma \in \mathcal{M}_k(N, \chi^\sigma)$.*

Proof. This is easy using the definition of $\mathcal{M}_k(N, \chi)$. □

Proposition 2.6.4. *For all $n \geq 1$, all $\sigma \in G$ and $f \in \mathcal{M}_k(N, \chi)$*

$$T_n(f^\sigma) = (T_n(f))^\sigma.$$

Proof. For any $m \geq 0$

$$(a_m(T_n f))^\sigma = \left(\sum_{d|(n,m)} \chi(d) d^{k-1} a_{nm/d^2}(f) \right)^\sigma = \sum_{d|(n,m)} \chi(d)^\sigma d^{k-1} a_{nm/d^2}(f)^\sigma$$
$$= \sum_{d|(n,m)} \chi(d)^\sigma d^{k-1} a_{nm/d^2}(f^\sigma) = a_m(T_n(f^\sigma)),$$

where in the last equality we use the above lemma. □

Theorem 2.6.5. *Let k and N be integers and χ a Dirichlet character modulo N. Suppose that $f = \sum_{n=0}^{\infty} a_n(f) q^n$ is a Hecke form in $\mathcal{M}_k(N, \chi)$ (resp. $\mathcal{S}_k(N, \chi)$). Then, there is a number field whose ring of integers contain the coefficients $\{a_n\}_{n\geq 0}$.*

Proof. Let $\mathcal{B} = \{f_1, \ldots, f_d\}$ a basis of Hecke forms in $\mathcal{M}_k(N, \chi)$. We can suppose $f \in \mathcal{B}$. For all $n \geq 1$ and all $\sigma \in G$,

$$a_n(f^\sigma) f^\sigma = T_n(f^\sigma) = (T_n(f))^\sigma = (a_n(f) f)^\sigma = a_n(f)^\sigma f^\sigma,$$

then $(a_n(f))^\sigma = a_n(f^\sigma)$, for $n \geq 1$, but if we take $K \subset \mathbb{C}$ to be the field generated by $\{a_n\}_{n\geq 1}$ and $\tau \in \mathrm{Aut}_K(\mathbb{C})$ then $a_0 - a_0^\tau = f - f^\tau$ is a constant modular form, so is 0, and then $a_0 = a_0^\tau$, thus a_0 can be obtained by a linear combination of the other coefficients. Now, $f, f^\sigma \in \mathcal{B}$, so for all σ there exists $i \in \{1, \ldots, d\}$ such that $f^\sigma = f_i$. Therefore, the action of G in $\mathcal{B}$ factorises through a finite group H. By Galois theory, there is a Galois finite extension $L/\mathbb{Q}$ such that the orbit of H and $\mathrm{Gal}(L/\mathbb{Q})$ are the same.

Thus, for all $\sigma \in \mathrm{Aut}_L(\mathbb{C})$

$$\sigma(a_m(f)) = a_m(f),$$

i.e., $a_m(f) \in L$.

Now take V to be the lattice generated by the basis of Hecke forms $\{f_1, \ldots, f_d\}$ over the ring of integers of L which we will denote by $\mathcal{O}$. Then, for any n $T_n V \subseteq V$, so we can express the action of T_n over V as a $d \times d$ matrix M_n. The characteristic polynomial of M_n is a monic polynomial over $\mathcal{O}$ which has as roots the eigenvalues of $\{f_1, \ldots, f_d\}$, and they are exactly $\{a_n(f_1), \cdots, a_n(f_d)\}$, which proves that the basis has Fourier coefficients in $\mathcal{O}$.

□

Remark 2.6.6. *We are using strongly that $a_1(f_i) = 1$ for all i, otherwise, $T_n f = a_1 a_n$. So if we multiply by a transcendent factor a Hecke form we would get an eigenform not normalised and the proof would not apply.*

2.7 Oldforms and newforms

Definition 2.7.1. *Let N, k, d be integers such that $d \mid N$. Let $\gamma_d = \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix}$. Consider the map*

$$\begin{array}{rccc} i_d : & (\mathcal{S}_k(\Gamma_1(N/d)))^2 & \longrightarrow & \mathcal{S}_k(\Gamma_1(N)) \\ & (f,g) & \longmapsto & f + g|[\gamma]_k. \end{array}$$

We define the subspace of ***oldforms at level*** *N by*

$$\mathcal{S}_k(\Gamma_1(N))^{\text{old}} = \sum_{p|N \ \ prime} i_p((\mathcal{S}_k(\Gamma_1(N/p)))^2)$$

and the subspace of ***newforms at level*** *N by*

$$\mathcal{S}_k(\Gamma_1(N))^{\text{new}} = (\mathcal{S}_k(\Gamma_1(N))^{\text{old}})^{\perp},$$

i.e., the orthogonal with respect to the Petersson inner product.

The idea of these forms is that if $M|N$ we have that $\Gamma_1(N) \subseteq \Gamma_1(M)$ because of obvious reasons. Hence, in the definition of modular form there is only one condition concerning the group, and this is

$$f|[\gamma](z) = f(z), \quad \forall \gamma \in \Gamma, \forall z \in \mathbb{H},$$

but if it is true for all $\gamma \in \Gamma_1(M)$ since $\Gamma_1(N) \subseteq \Gamma_1(M)$ we have that it is also true for the elements in $\Gamma_1(N)$ and then

$$\mathcal{M}_k(\Gamma_1(M)) \subseteq \mathcal{M}_k(\Gamma_1(N)).$$

Intuitively this is understandable, putting a bigger group means having *more* conditions, and this means *less* modular forms.

We call these forms oldforms because they belonged to the first space and are not *new*.

Proposition 2.7.2 (Cf.[DS05] Proposition 5.6.2)**.** *The spaces $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ and $\mathcal{S}_k(\Gamma_1(N))^{\text{old}}$ are stable under the Hecke operators T_n and $\langle n \rangle$ for any n.*

This proposition will be important in the last chapters because it will ensure the existence of congruences between oldforms and newforms. The proof is not given.

Corollary 2.7.3. *The spaces $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ and $\mathcal{S}_k(\Gamma_1(N))^{\text{old}}$ have orthogonal bases of eigenforms of the operators T_n and $\langle n \rangle$ such that $\gcd(n, N) = 1$.*

Definition 2.7.4. *A* ***newform*** *is a normalised Hecke form belonging to $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$.*

Chapter 3

Classical congruences between modular forms and generalisations

This chapter tries to introduce the first methods in order to get congruences between modular forms. The title of this chapter contains the word classic, and even if there are some classic results in this chapter, some others are recent. However, the methods used in the proofs of all the theorems in this chapter use only basic Number Theory and the ones that need more advanced methods are proved entirely in this section or in the preceding ones.

3.1 Definition

Definition 3.1.1. *Let f, g be two modular forms of the same fixed level and weight. As we have seen in chapter 2, we can assume they have Fourier coefficients in an algebraic integral ring $\mathcal{O}$, put $\{a_n(f)\}_n, \{a_n(g)\}_n$ respectively. Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}$, we say that f and g are congruent modulo $\mathfrak{p}$ if and only if, for all $n \geq 0$ $a_n(f) - a_n(g) \in \mathfrak{p}$. We denote it by*

$$a_n(f) \equiv a_n(g) \pmod{\mathfrak{p}}$$

and

$$f \equiv g \pmod{\mathfrak{p}}.$$

3.2 Congruences of Eisenstein and cusps forms on level N=1

This is the very first case, in 1916 Ramanujan showed that $\tau(n) \equiv \sigma_{11}(n) \pmod{691}$ which is remarkable. In the language used in this thesis we prefer to state that

$$G_{12}(z) \equiv \Delta(z) \pmod{691}.$$

In this section we will reproduce the generalising results in [DG96]. So we are in the case of modular forms with respect to the full modular group with a fixed weight k. Recall

$$G_k(z) = -\frac{B_k}{2k} + \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n, \quad q = e^{2\pi i z}.$$

Theorem 3.2.1. *Let N be the numerator of the reduced fraction $\frac{B_k}{2k}$. For every $k \geq 12$ there exists $f \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z})) \setminus \{0\}$ such that*

$$f \equiv G_k \pmod{N}.$$

Proof. Consider the normalised Eisenstein series E_4, E_6. Consider the Euclidean division $k = 6q + r$, since k is even, r is either $0, 2$ or 4, so define

$$g := \begin{cases} E_6^q, & \text{if } r = 0, \\ E_6^{q-1} E_4^2, & \text{if } r = 2, \\ E_6^q E_4, & \text{if } r = 4 \end{cases}$$

g is always a modular form of weight k which has its first Fourier coefficient equal to 1.

Define $f = G_k + \frac{B_k}{2k} g$, it is a cusp form and

$$G_k \equiv f \pmod{N}.$$

For $k \geq 12$, $N > 1$, and $a_1(G_k) = 1$, then $f \neq 0$. □

An immediate corollary is:

Corollary 3.2.2. *If $\dim_{\mathbb{C}} \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z})) = 1$, and f is a non-zero cusp form with $a_1(f) = 1$, then*

$$f \equiv G_k \pmod{N}.$$

Example 3.2.3 (Ramanujan)**.** *If $k = 12$, we only have one cusp form which is $\Delta(z)$ and*

$$\frac{B_{12}}{24} = -\frac{691}{65520},$$

so $N = 691$ *and*

$$G_{12}(z) \equiv \Delta(z) \pmod{691}.$$

This last congruence is obviously equivalent to the congruences between the Ramanujan's τ*-function and* σ_{11}*.*

When our cusp space has a larger dimension, we can find some congruences using a basis of Hecke forms. Before formulate a theorem, we need the following lemma.

Lemma 3.2.4. *Let* $\{f_1, \dots, f_r\}$ *be a basis of the cusp space of Hecke forms. Let* K *be the number field containing all Fourier coefficients of* $f_1, \dots, f_r$*. Let* $\mathcal{O}_K$ *be the ring of integers of* K*. Let* N *be as above and* $\mathfrak{p}$ *a prime ideal of* $\mathcal{O}_K$ *dividing* N*. Assume that there exist* $\beta_0, \dots, \beta_r \in \mathcal{O}_K$ *and* $m \in \mathbb{N}$ *such that* $\beta_0 \not\equiv 0 \pmod{\mathfrak{p}^m}$ *and*

$$\beta_1 f_1 + \cdots + \beta_r f_r \equiv \beta_0 G_k \pmod{\mathfrak{p}^m},$$

then there exists i *such that*

$$f_i \equiv G_k \pmod{\mathfrak{p}}.$$

Proof. Choose a minimal subset of $\{f_1, \dots, f_r\}$ with the property described in the hypothesis. Say $\{f_1, \dots, f_l\}$ up to reordering. Then for some m

$$\beta_1 f_1 + \cdots + \beta_l f_l \equiv \beta_0 G_k \pmod{\mathfrak{p}^m}. \tag{3.1}$$

Apply T_n to the equality, then

$$\beta_1 a_n(f_1) f_1 + \cdots + \beta_l a_n(f_l) f_l \equiv \beta_0 a_n(G_k) G_k \pmod{\mathfrak{p}^m}. \tag{3.2}$$

Using both equalities, we cancel f_1 making the difference of equation 3.2 and $a_n(f_1)$ times equation 3.1. We get

$$\beta_2(a_n(f_2) - a_n f_1) f_2 + \cdots + \beta_l(a_n(f_l) - a_n(f_1)) f_l \equiv \beta_0(a_n(G_k) - a_n(f_1)) G_k \pmod{\mathfrak{p}^m}.$$

This contradicts the minimality of the chosen set unless $\beta_0(a_n(G_k) - a_n(f_1)) \equiv 0 \pmod{\mathfrak{p}^m}$, so,$a_n(G_k) - a_n(f_1) \equiv 0 \pmod{\mathfrak{p}}$. Then, doing this with all n we deduce that

$$f_1 \equiv G_k \pmod{\mathfrak{p}},$$

and $l = 1$. □

Theorem 3.2.5. *With the above notation, there exist* i *such that*

$$f_i \equiv G_k \pmod{\mathfrak{p}}.$$

Proof. By the first theorem in this section, there is a cusp form f congruent to G_k modulo N. We can express

$$f = \lambda_1 f_1 + \cdots + \lambda_r f_r,$$

for elements $\lambda_i \in K$. Choose $\lambda \in \mathcal{O}_K$ such that $\lambda\lambda_i \in \mathcal{O}_K$ for all i, denote $\beta_i = \lambda\lambda_i$. We have

$$\beta_1 f_1 + \cdots + \beta_r f_r \equiv \lambda G_k \pmod{\lambda N}.$$

For $\mathfrak{p}$ prime dividing N, take $m = \mathrm{ord}_{\mathfrak{p}}(\lambda) + 1$. Then, the above congruence is equivalent to

$$\beta_1 f_1 + \ldots \beta_r f_r \equiv \lambda G_k \pmod{\mathfrak{p}^m}.$$

By definition, $\lambda \not\equiv 0 \pmod{\mathfrak{p}^m}$. So by the lemma, there is i such that

$$f_i \equiv G_k \pmod{\mathfrak{p}}.$$

□

This is an important result, because links the Eisenstein series with the cusps of the same weight, providing congruences for certain prime ideals.

3.3 Sturm's approach

In [Stu84], Sturm gave a sufficient condition for congruences between modular forms. In a computational level it is very convenient to have a bound on the number of coefficients we need to compare to determine that two modular forms are congruent modulo some prime. Sturm proved a very nice Theorem which gives a bound.

Definition 3.3.1. *Let K be a number field and $\mathcal{O}_K$ its ring of integers. Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. If $f \in \mathcal{O}_K[[q]]$, where $q = e^{2\pi i z}$, we define the order of f by*

$$\mathrm{ord}_{\mathfrak{p}}(f) = \inf_n \{a_n(f) : \mathfrak{p} \nmid a_n(f)\}.$$

Lemma 3.3.2. *For f, g two modular forms with respect to Γ with coefficients in $\mathcal{O}$ and $\mathfrak{p}$ a prime ideal we have*

$$\mathrm{ord}_{\mathfrak{p}}(fg) = \mathrm{ord}_{\mathfrak{p}}(f) + \mathrm{ord}_{\mathfrak{p}}(g).$$

Proof. Let N be the level and let $f = \sum_{n\geq 0} a_n(f) q_N^n$, $g = \sum_{n\geq 0} a_n(g) q_N^n$ be the Fourier series of f and g. Then,

$$fg = \sum_{n\geq 0} a_n(fg) q_N^n,$$

with

$$a_n(fg) = \sum_{i+j=n} a_i(f) a_j(g).$$

Assume that $\mathrm{ord}_{\mathfrak{p}}(f) = r, \mathrm{ord}_{\mathfrak{p}}(g) = s$ are finite, then

$$a_{r+s}(fg) \equiv a_r(f)a_s(g) \not\equiv 0 \pmod{\mathfrak{p}}$$

and if $n < r + s$

$$a_n(fg) = \sum_{i+j=n} a_i(f)a_j(g) \equiv 0 \pmod{\mathfrak{p}}.$$

So $\mathrm{ord}_{\mathfrak{p}}(fg) = r+s$ if r, s are finite, but note that if one of these is infinite, for example f, this means that all its coefficients are divisible by $\mathfrak{p}$ and so is $\sum_{i+j=n} a_i(f)a_j(g)$. □

Theorem 3.3.3. *With the above notations, let $f, g \in \mathcal{M}_k(\Gamma)$ with Fourier coefficients in $\mathcal{O}_K$. If*

$$\mathrm{ord}_{\mathfrak{p}}(f - g) > \frac{k[\mathrm{SL}_2(\mathbb{Z}) : \Gamma]}{12},$$

then

$$ord_{\mathfrak{p}}(f - g) = \infty,$$

i.e.,

$$f \equiv g \pmod{\mathfrak{p}}.$$

In order to prove the Theorem we need to prove the following lemma, we follow Ram Murty [Mur97].

Lemma 3.3.4. *Let $\Phi \in \mathcal{M}_{12k}(\mathrm{SL}_2(\mathbb{Z}))$ satisfying $\mathrm{ord}_{\mathfrak{p}}(\Phi) > k$. Then*

$$\frac{\Phi}{\Delta^k} \in \mathfrak{p}[j].$$

I.e., it can be written as a polynomial in j whose coefficients are all divisible by $\mathfrak{p}$.

Proof. We will prove it by induction on k. For $k = 1$, since $\Phi \in \mathcal{M}_{12}(\mathrm{SL}_2(\mathbb{Z})) = \langle E_{12}, \Delta \rangle$ and $E_{12} = E_4^3$, so $\Phi = \lambda E_4^3 + \mu\Delta$. Recall that $j = E_4^3/\Delta$, so we have

$$\frac{\Phi}{\Delta} = \lambda j + \mu.$$

And from $\mathrm{ord}_{\mathfrak{p}}(\Phi) > 1 \implies \mathrm{ord}_{\mathfrak{p}}(\Phi/\Delta) > 0$. This means that if $\Phi/\Delta = \sum_{n \geq -1} a_n q^n$ we have that $\mathfrak{p} | c_{-1}, c_0$ which means that $\mathfrak{p}$ divides λ and then μ.

Now let $k > 1$ and i, j such that $12k = 4i + 6j$. This implies that for some $c\mathcal{O}$

$$\Phi - cE_4^i E_6^j$$

is a cusp form, so we can write it

$$\Phi = cE_4^i E_6^j + \Delta f_1,$$

where $f_1 \in \mathcal{M}_{12(k-1)}(\mathrm{SL}_2(\mathbb{Z}))$. Since $cE_4^iE_6^j = c(1 + a_1q + \cdots)$ and $\mathfrak{p}|a_n(\Phi)$ for $n \leq k$, provided that $k > 1$ we deduce $\mathfrak{p}|c$ and hence $\mathfrak{p}$ divides the first $k-1$ coefficients of f_1. By induction hypothesis, $f_1/\Delta^{k-1} \in \mathfrak{p}[j]$. Now,

$$\frac{\Phi}{\Delta^k} = c\frac{E_4^iE_6^j}{\Delta^k} + \frac{f_1}{\Delta^{k-1}}.$$

From $12k = 4i + 6j$ follows that $i = 3i_0, j = 2j_0$ $i_0, j_0 \in \mathbb{N}$, now $E_4^3/\Delta = j$, $E_6^2/\Delta = j - 1728$ completes the proof of the lemma. □

Proof of Theorem 3.3.3. We begin assuming $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. Let $h = f - g$ which is automorphic and has a single pole at infinity. By hypothesis $\mathrm{ord}_{\mathfrak{p}}(h^{12}) > k$, then

$$h^{12}\Delta(z)^{-k} = \sum_{n \geq -k} c_n q^n$$

with $c_n \in \mathcal{O}_K$ and $\mathfrak{p}|c_n$ for $n \leq 0$. Then $h^{12}\Delta^{-k} \in \mathfrak{p}[j]$ of degree at most k, so $h^{12} \in \mathfrak{p}[j]\Delta^k$, which implies $\mathrm{ord}_{\mathfrak{p}}(h^{12}) = \infty$ and then $\mathrm{ord}_{\mathfrak{p}}(h) = \mathrm{ord}_{\mathfrak{p}}(f - g) = \infty$.

Assume now Γ arbitrary and define as before $h = f - g$, we may assume $12|k$ (if it is not the case , without loss of generality replace h by h^{12}). Consider $h\Delta^{-k/12}$ and applying theorem 6.6 of [Shi71] we conclude that for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ $h|[\gamma]_k \in \mathcal{M}(\Gamma(N))$ with coefficients in $K(\zeta)$, with $\zeta^N = 1$ primitive. Let $L|K(\zeta)$ be a finite extension such that $\mathfrak{p}\mathcal{O}_L$ is principal and let $\mathfrak{P}$ be a prime dividing $\mathfrak{p}$. Then, for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, by the Chinese Reminder theorem there is $\alpha_\gamma \in \mathcal{O}_L$ non-trivial such that $\alpha_\gamma h|[\gamma]_k$ has Fourier coefficients in $\mathcal{O}_L$. And, $v_{\mathfrak{P}}(\alpha_\gamma h|[\gamma]_k)$ is finite.

Write

$$\mathrm{SL}_2(\mathbb{Z}) = \bigcup_{i=1}^{m} \Gamma\gamma(i)$$

where $\gamma(i) = \mathrm{Id}$. Let

$$H = h(\prod_{i=1}^{m} \alpha_{\gamma_i} h|[\gamma_i]_k).$$

Then $H \in \mathcal{M}_{km}(\mathrm{SL}_2(\mathbb{Z}))$ and $v_{\mathfrak{P}}(H) \geq v_{\mathfrak{P}}(h) = v_{\mathfrak{p}}(h) > km/12$. Hence, by the case $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ we have $v_{\mathfrak{P}}(H) = \infty$. So

$$v_{\mathfrak{P}}(h) + \sum_{i=1}^{m} v_{\mathfrak{P}}(\alpha_{\gamma_i} h|[\gamma_i]_k) = \infty,$$

which implies $v_{\mathfrak{P}}h = v_{\mathfrak{p}}h = \infty$. □

Remark 3.3.5. *1. If $\Gamma = \Gamma(N)$, then the bound is*

$$\mathrm{ord}_{\mathfrak{p}}(f - g) > \frac{kN^3 \prod_{p|N}(\frac{p^2-1}{p^2})}{12}.$$

2. *If* $\Gamma = \Gamma_0(N)$, *then the bound is*

$$\mathrm{ord}_{\mathfrak{p}}(f-g) > \frac{kN\prod_{p|N}(\frac{p-1}{p})}{12}.$$

3. *If* $\Gamma = \Gamma_1(N)$, *then the bound is*

$$\mathrm{ord}_{\mathfrak{p}}(f-g) > \frac{kN\varphi(N)\prod_{p|N}(\frac{p-1}{p})}{12}.$$

Example 3.3.6. *In the case* $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ *the bound is much simpler. For example, the case* $\Delta(z) \equiv G_{12}(z) \pmod{691}$, *we only need to see that*

$$\mathrm{ord}_{691}(\Delta(z) - G_{12}(z)) > 1,$$

so

$$\tau(0) \equiv 0 \pmod{691}, \tau(1) \equiv 1 \pmod{691}$$

and

$$\tau(2) \equiv 1 + 2^{11} \pmod{691}.$$

In this case it is simple computation, but this approach is giving no information about new congruences nor existence of them.

The result can be improved when we treat newforms.

Theorem 3.3.7 (Cf. [Stu84] Theorem 2.)**.** *Let* $f, g \in \mathcal{S}_k(\Gamma_0(N), \chi)^{\mathrm{new}}$ *for some Dirichlet character*

$$\chi : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{C}^*.$$

Assume that the Fourier expansions of f *and* g *have coefficients in* $\mathcal{O}_K$ *as before. Let* $\{p_1, \ldots, p_r\}$ *be a subset of the prime divisors of* N. *If*

1. $\mathrm{ord}_{\mathfrak{p}}(f-g) > \frac{k\sigma_1(N)}{2^r \cdot 12}$,
2. $a_{p_i}(f) = a_{p_i}(g)$, *for* $i = 1, \ldots, r$.

Then,

$$f \equiv g \pmod{\mathfrak{p}}.$$

If $\chi = \mathbb{1}$ *and* $\mathfrak{p}|2$,

$$\mathrm{ord}_{\mathfrak{p}}(f-g) > \frac{k\sigma_1(N)}{12\sigma_0(N)} \implies f \equiv g \pmod{\mathfrak{p}}.$$

3.4 Generalising Ramanujan congruences

This is a result concerning Ramanujan congruences found in [DF14].

Theorem 3.4.1. *Let p, ℓ be prime numbers and $k \geq 4$ an even integer. Assume $\ell > 3$ and $v_\ell((p^k-1)(B_k/2k)) > 0$. Then there is a normalised Hecke form (for all T_q $q \neq p$ prime) $f = \sum_{n=1}^{\infty} a_n(f)q^n \in \mathcal{S}_k(\Gamma_0(p))$, and $\mathfrak{l}|\ell$ in the ring of integers $\mathcal{O}$ of the field $\mathbb{Q}_f = \mathbb{Q}(\{a_n(f)\})$, such that*

$$a_q(f) \equiv 1 + q^{k-1} \pmod{\mathfrak{l}}.$$

Proof. Define $h(z) = E_k(z) - E_k(pz)$ which has its first Fourier coefficient equal to 0. We need to know what happens in 0. We use the W_p operator which switches 0 and ∞

$$\begin{aligned} W_p(h(z)) =& W_p(E_k(z)) - W_p(E_k(pz)) = z^{-k}E_k(-1/pz) - z^{-k}E_k(-p/pz) \\ =& z^{-k}p^k z^k E_k(pz) - z^{-k}z^k E_k(z) = p^k E_k(z) - E_k(z). \end{aligned}$$

This form at infinity is $(p^k-1)\frac{-B_k}{2k} \equiv 0 \pmod{\ell}$. So the reduction modulo ℓ of h is a mod ℓ cusp form. By the surjectivity of the map

$$\mathcal{S}_k(\Gamma_0(p), \mathbb{Z}_\ell) \longrightarrow \mathcal{S}_k(\Gamma_0(p), \mathbb{F}_\ell),$$

(using $\ell > 3$) there is $g \in \mathcal{S}_k(\Gamma_0(p), \mathcal{O}'_\mathfrak{l})$ where $\mathcal{O}'_\mathfrak{l}$ is the ring of integers of a finite extension of $\mathbb{Q}_\ell$. Let $\mathbb{F}$ be the residue field of $\mathcal{O}'_\mathfrak{l}$, the reduction of g in this field is a common eigenvector for T_q for $q \neq p$ prime with eigenvalue $1 + q^{k-1}$ and by a Lemma of Deligne and Serre, there is $f' \in \mathcal{S}_k(\Gamma_0(p), \mathcal{O}_\mathfrak{l})$ a Hecke form in a finite extension of the above extenion with eigenvalues

$$a_q \equiv 1 + q^{k-1} \pmod{\mathfrak{l}}.$$

This form arises from a form with coefficients in $\mathbb{Q}_f$. □

3.5 Congruences using $X_0(N)$

We give here a very interesting result using the geometry in $X_0(N)$ applying Riemann-Roch Theorem. This result is due to [Mur97]. These are not congruences in a strict sense, but congruences with the prime at infinity, i.e., we find equalities between modular forms of differents levels and weights. Using these results we will find proper congruences.

Theorem 3.5.1. *Assume that $f \in \mathcal{M}_k(\Gamma_0(N_1))$ and $h \in \mathcal{M}_k(\Gamma_0(N_2))$, let $N = \operatorname{lcm}(N_1, N_2)$ assume that*

$$\operatorname{ord}_\infty(f - h) > \frac{k}{2}(2g - 1),$$

where g is the genus of $X_0(N)$. Then $f = h$.

Proof. Suppose $f \neq h$, then k is even because in $-Id \in \Gamma_0(N)$ for all N. Take $\omega = (f-h)(dz)^{k/2}$ which is a holomorphic $k/2$ form on $X_0(N)$. Its degree is $\frac{k}{2}(2g-2)$. Now, the hypothesis implies that $\mathrm{ord}_{i\infty}(\omega) \geq \frac{k}{2}(2g-1) - \frac{k}{2}$. Hence

$$\frac{k}{2}(2g-2) = \deg(\omega) \geq \mathrm{ord}_{i\infty}(\omega) \geq \frac{k}{2}(2g-1) + \frac{k}{2},$$

and this is a contradiction, because the genus is positive.

In particular, ω is the zero form and hence $f = h$. □

Theorem 3.5.2. *Let us assume now that f has weight k_1 and g has weight k_2, then if*

$$\mathrm{ord}_{\infty}(f-g) > k_1 k_2 (g-1)$$

we have $f = g$.

Proof. The proof is very similar, assume $f \neq h$ and that k_1, k_2 are even. Define $\omega = (f^{k_2} - h^{k_1})(dz)^{k_1 k_2/2}$. Then

$$\mathrm{ord}_{i\infty}(\omega) \geq \frac{k_1 k_2}{2}(2g-1) + \frac{k_1 k_2}{2} = k_1 k_2 g.$$

Then,

$$\frac{k_1 k_2}{2} = \deg(\omega) \geq \mathrm{ord}_{i\infty}(\omega) \geq k_1 k_2 g$$

which is a contradiction and implies $\omega = 0$ and $f = h$. □

Now we will apply the same method to find proper congruences using Riemann-Roch Theorem (valid in characteristic p if $(N, p) = 1$).

Theorem 3.5.3. *Let f, h be cusps forms of weight k and level N with coefficients lying in some ring of integers of some number field K.Let $\mathfrak{p}$ be a prime ideal such that $N \notin \mathfrak{p}$, then if*

$$\mathrm{ord}_{\mathfrak{p}}(f-h) > \frac{k}{2}(2g-1)$$

we have that $f \equiv h \pmod{\mathfrak{p}}$.

Proof. The proof is exactly the same as before because we can apply Rieman-Roch and for all N coprime to $\mathfrak{p}$ there is a good reduction of $X_0(N)$. □

In the same way we have

Theorem 3.5.4. *Let us assume now that f has weight k_1 and level N_1 and h has weight k_2 and level N_2. Assume that f, g have integral coefficients for some number field K then if $N \notin \mathfrak{p}$ and if*

$$\mathrm{ord}_{\mathfrak{p}}(f-g) > k_1 k_2 (g-1)$$

we have $f \equiv h \pmod{\mathfrak{p}}$.

3.6 Congruences modulo $\mathfrak{p}^m$

There is a nice generalisation due to Chen, Kiming and Rasmussen [IC10] in which they generalise the condition of Sturm for powers of primes.

Even if we will not go into the theory of congruences modulo prime powers, this result is easy to understand and gives an insight of the strength of these computational methods.

Proposition 3.6.1. *Let $f, g \in \mathcal{M}_k(\Gamma_1(N))$ and let $\mathcal{O}$ a ring containing their Fourier coefficients. Then, if*

$$\operatorname{ord}_{\mathfrak{p}^m}(f-g) > \frac{k[\mathrm{SL}_2(\mathbb{Z}):\Gamma_1(N)]}{12}$$

we have that

$$f \equiv g \pmod{\mathfrak{p}^m}.$$

Proof. We will prove it by induction on m. The initial case follows from the theorem by Sturm (3.3.3) but not immediately, we should prove it for $\mathcal{O}_\mathfrak{p}$ the localised ring. If $h = f - g$ has Fourier coefficients in $\mathcal{O}_\mathfrak{p}$ then there is $\lambda \in \mathcal{O} \setminus \mathfrak{p}$ such that λh has coefficients in $\mathcal{O}$, hence

$$\operatorname{ord}_\mathfrak{p}(h) = \operatorname{ord}_\mathfrak{p}(\lambda h) > \frac{k[\mathrm{SL}_2(\mathbb{Z}):\Gamma_1(N)]}{12}$$

by Sturm we have

$$\operatorname{ord}_\mathfrak{p}(h) = \operatorname{ord}_\mathfrak{p}(\lambda h) = \infty.$$

Assume now that this result is true for powers $\mathfrak{p}^n$ with $n < m$ and $m > 1$ for coefficients in $\mathcal{O}_\mathfrak{p}$.

Let $h = f - g$ and assume $\operatorname{ord}_{\mathfrak{p}^m}(h) > \frac{k[\mathrm{SL}_2(\mathbb{Z}):\Gamma_1(N)]}{12}$. In particular, this implies that

$$\operatorname{ord}_{\mathfrak{p}^{m-1}}(h) > \frac{k[\mathrm{SL}_2(\mathbb{Z}):\Gamma_1(N)]}{12},$$

by inductive hypothesis we then know that

$$\operatorname{ord}_{\mathfrak{p}^{m-1}}(h) = \infty.$$

Let $\pi \in \mathfrak{p} \setminus \mathfrak{p}^2$ a uniformiser. Now define

$$H = \frac{1}{\pi^{m-1}} h$$

and of course $H \in \mathcal{M}_k(\Gamma_1(N))$ which has coefficients in $\mathcal{O}_\mathfrak{p}$, because $\operatorname{ord}_{\mathfrak{p}^{m-1}}(h) = \infty$ implies that all its Fourier coefficients are divisible by $\mathfrak{p}^{m-1}$, in particular by π^{m-1}.

Since $\mathrm{ord}_{\mathfrak{p}^m}(h) > \frac{k[\mathrm{SL}_2(\mathbb{Z}):\Gamma_1(N)]}{12}$ we have that $\mathrm{ord}_{\mathfrak{p}}(H) > \frac{k[\mathrm{SL}_2(\mathbb{Z}):\Gamma_1(N)]}{12}$, so by Sturm

$$\mathrm{ord}_{\mathfrak{p}}(H) = \infty,$$

and hence

$$\mathrm{ord}_{\mathfrak{p}^m}(h) = \mathrm{ord}_{\mathfrak{p}^m}(f-g) = \infty,$$

i.e.

$$f \equiv g \pmod{\mathfrak{p}^m}.$$

□

Chapter 4

Congruences related to ℓ-adic representations

4.1 ℓ-adic representations

In this section ℓ will denote a prime number and $\overline{\mathbb{Q}}$ will denote a fixed algebraic closure of $\mathbb{Q}$.

Definition 4.1.1. *The **absolute Galois group** is defined as the group of automorphisms of $\overline{\mathbb{Q}}$. It is denoted as*

$$G_{\mathbb{Q}} = \mathrm{Aut}(\overline{\mathbb{Q}}).$$

Proposition 4.1.2. *Let $\{K_i\}_i$ be the set of all Galois extensions of $\mathbb{Q}$, then*

$$G_{\mathbb{Q}} = \varprojlim_i \{\mathrm{Gal}(K_i/\mathbb{Q})\}.$$

Proof. The map

$$\begin{array}{rccc} \psi_i : & G_{\mathbb{Q}} & \longrightarrow & \mathrm{Gal}(K_i/\mathbb{Q}) \\ & \sigma & \longmapsto & \sigma\,|_{K_i} \end{array}$$

surjects, and if $K_i \subseteq K_j$ then

$$\sigma\,|_{K_i} = \sigma\,|_{K_j}|_{K_i}$$

in other words $\psi_i = \psi_j \circ \pi_{j,i}$, where

$$\begin{array}{rccc} \pi_{j,i} : & \mathrm{Gal}(K_j/\mathbb{Q}) & \longrightarrow & \mathrm{Gal}(K_i/\mathbb{Q}) \\ & \tau & \longmapsto & \tau\,|_{K_i} \end{array}$$

Conversely, every compatible chain of automorphisms of Galois fields define an automorphism of $\overline{\mathbb{Q}}$.

□

In order to study some properties of the absolute Galois group we can define a topology that will give us many information.

Definition 4.1.3. *The* ***Krull topology*** *is the topology generated by the sets*

$$U_\sigma(K) = \{\sigma\tau : \ \tau\,|_K = 1\},$$

where $K/\mathbb{Q}$ is a Galois extension and $\sigma \in G_\mathbb{Q}$.

Observation 4.1.4. *We will denote $U_1(K)$ by $U(K)$.*

Definition 4.1.5. *Let $r \in \mathbb{Z}_{>0}$. A d-**dimensional** ℓ-**adic Galois representation** is a continuous homomorphism*

$$\rho : G_\mathbb{Q} \longrightarrow \mathrm{GL}(d, L)$$

where L is a finite extension of $\mathbb{Q}_\ell$.

If we have two such representations ρ, ρ', we say they are equivalent if, and only if there is a matrix $m \in \mathrm{GL}(d, L)$ such that $\rho(\sigma) = m^{-1}\rho(\sigma)m$ for all $\sigma \in G_\mathbb{Q}$. We write it $\rho \sim \rho'$.

Observation 4.1.6. *In the above definition we implicitly use a topology in $\mathrm{GL}(d, L)$ which is the topology induced by $\mathbb{Q}_\ell$ and the relations defining $\mathrm{GL}(d, L)$.*

In order to show some properties of these new objects we need more knowledge of the arithmetic of $G_\mathbb{Q}$.

Definition 4.1.7. *Let $\overline{\mathbb{F}}_p$ be a fixed algebraic closure of $\mathbb{F}_p$. The* ***absolute Galois group of*** *$\mathbb{F}_p$ is*

$$G_{\mathbb{F}_p} = \mathrm{Aut}(\overline{\mathbb{F}_p}).$$

Definition 4.1.8. *Let K be a number field and let L/K be a finite Galois extension. Let $\mathfrak{P}$ be a prime ideal of the integers of L and $\mathfrak{p}$ a prime ideal of the integers of K such that $\mathfrak{P}$ divides $\mathfrak{p}$. Consider $L_\mathfrak{P}$ and $K_\mathfrak{p}$ the corresponding completions of L and K. We define the decomposition group to be*

$$D_\mathfrak{P} = \mathrm{Gal}(L_\mathfrak{P}/K_\mathfrak{p}).$$

It is equivalent to

$$D_\mathfrak{P} = \{g \in \mathrm{Gal}(L/K) : \ g(\mathfrak{P}) = \mathfrak{P}\}.$$

Definition 4.1.9. *With the same notation, call $l_\mathfrak{P}$ and $k_\mathfrak{p}$ the corresponding residue fields. We can map homomorphically $D_\mathfrak{P}$ onto $\mathrm{Gal}(l_\mathfrak{P}/k_\mathfrak{p})$. We call the inertia group to its kernel. It is denoted by $I_\mathfrak{P}$. It is equivalent to*

$$I_\mathfrak{P} = \{g \in D_\mathfrak{P} : \ x^g \equiv x \pmod{\mathfrak{P}} \text{ for all } x \in \mathcal{O}\}.$$

Definition 4.1.10. *An **absolute Frobenius element over** p is any pre-image* $\mathrm{Frob}_{\mathfrak{p}} \in D_{\mathfrak{p}}$ *of the Frobenius automorphism of* $G_{\mathbb{F}_p}$.

Lemma 4.1.11. *$D_{\mathfrak{P}}/I_{\mathfrak{P}}$ is a cyclic group and the equivalence class of a generator is the Frobenius element.*

Proof. This follows from the following exact sequence

$$1 \longrightarrow I_{\mathfrak{P}} \longrightarrow D_{\mathfrak{P}} \longrightarrow \mathrm{Gal}(\mathbb{F}_{\mathfrak{P}}/\mathbb{F}_{\mathfrak{p}}) \longrightarrow 1$$

Then

$$D_{\mathfrak{P}}/I_{\mathfrak{P}} \simeq \mathrm{Gal}(\mathbb{F}_{\mathfrak{P}}/\mathbb{F}_{\mathfrak{p}}) = \langle \varphi \rangle$$

where φ is the Frobenius automorphism. □

Definition 4.1.12. *Let $\rho : \mathrm{Gal}(\overline{K}/K) \longrightarrow \mathrm{GL}(n, \mathbb{Q}_\ell)$ be a n-dimensional ℓ-adic representation of $\mathrm{Gal}(\overline{K}/K)$, and let $\mathfrak{p}$ be a prime ideal of the integers of K. We say that ρ is **unramified** at $\mathfrak{p}$ if for any prime ideal $\mathfrak{P}$ of the integers of $\overline{K}$ dividing $\mathfrak{p}$ $\rho(I_{\mathfrak{P}}) = \{\mathrm{id}\}$.*

4.2 Representations attached to modular forms

Let ℓ be a prime number, K a maximal extension of $\mathbb{Q}$ such that the only prime ramified is ℓ, and K^{ab} a maximal sub-extension such that $\mathrm{Gal}(K^{ab}/\mathbb{Q})$ is abelian.

Proposition 4.2.1. *With the above notations,*

$$\mathrm{Gal}(K^{ab}) \simeq \mathbb{Z}_\ell^*.$$

This isomorphism induces a canonical character

$$\chi : \mathrm{Gal}(K/\mathbb{Q}) \longrightarrow \mathbb{Z}_\ell^*,$$

such that $\chi(\mathrm{Frob}(p)) = p$ for all prime different from ℓ.

Theorem 4.2.2 (Cf. [Del69])**.** *Let $f \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ with Fourier coefficients $\{a_n\}_n \subseteq \mathbb{Z}$, $a_1 = 1$ and assume it is a Hecke form. Then there is a continuous homomorphism*

$$\rho_f : \mathrm{Gal}(K/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\mathbb{Z}_\ell),$$

such that $\rho_f(\mathrm{Frob}(p))$ has characteristic polynomial

$$X^2 - a_p X + p^{k-1},$$

for each prime p different from ℓ.

Remark 4.2.3. *As an example, the modular form Δ (cf. section 1.1) satisfies the hypothesis since the Ramanujan τ function is multiplicative and takes values in $\mathbb{Z}$.*

4.3 Congruences on $\mathrm{SL}_2(\mathbb{Z})$ for ℓ

In this section we try to spell out the results that Swinnerton-Dyer presented in [SD73] which solved the problem of congruences on the full modular group for primes ℓ.

We will use primary the result conjectured by Serre and proved by Deligne in [Del69] which is Theorem 4.2.2.

This theorem states that for any $f \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ with Fourier coefficients in $\mathbb{Z}$, $a_1(f) = 1$ which is a Hecke form, we have a Galois ℓ-adic representation ρ_ℓ depending strongly on ℓ and f. The study of the possible images of ℓ-adic representations will provide information of the modular form modulo a prime.

Definition 4.3.1. *If f is as above, and ℓ is a prime number, if the image of ρ_ℓ does not contain $\mathrm{SL}_2(\mathbb{Z}_\ell)$ we say that ℓ is an **exceptional prime for** f.*

It may seem difficult to handle with such a condition. Next lemma reduces the problem for almost all primes.

Lemma 4.3.2 (Cf. [SD73] Lemma 1.)**.** *Assume $\ell > 3$ is a prime number, let G be a subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ closed in the ℓ-adic topology. Take $\overline{G}$ be the image of G by the projection on $\mathrm{GL}_2(\mathbb{F}_\ell)$. If $\mathrm{SL}_2(\mathbb{F}_\ell) \subseteq \overline{G}$, then $\mathrm{SL}_2(\mathbb{Z}_\ell) \subseteq G$.*

In order to understand the results, we need some background on group theory and a few lemmas that we will not prove.

Definition 4.3.3. *Let G be a subgroup of $\mathrm{GL}(\mathbb{F}_\ell)$, we say that G is a **Borel subgroup** if it is conjugate to the group of non-singular upper triangular matrices.*

Remark 4.3.4. *As $\mathrm{GL}(\mathbb{F}_\ell)$ acts on a 2-dimensional $\mathbb{F}_\ell$-vector space V, we can associate a one to one correspondence between Borel subgroups and one-dimensional spaces W of V (explicitly, the group having W as eigenspace).*

Definition 4.3.5. *Let G be a subgroup of $\mathrm{GL}(\mathbb{F}_\ell)$, we say that G is a **Cartan subgroup** if it is a maximal semi-simple (i.e. all its connected closed normal subgroups are trivial) abelian group.*

*A **split Cartan subgroup** is any subgroup conjugate to the group of non-singular diagonal matrices. It is, of course, a Cartan subgroup.*

*Take now V the 2-dimensional $\mathbb{F}_\ell$-vector space on which $\mathrm{GL}_2(\mathbb{F}_\ell)$ acts. Define $V_2 = V \otimes \mathbb{F}_{\ell^2}$. Let W be a one-dimensional subspace of V_2 not induced by a space of V. Take W' the conjugate of E over $\mathbb{F}_\ell$. We associate to W' the group formed by the elements of $\mathrm{GL}_2(\mathbb{F}_\ell)$ which have W' as eigenspace. If G is one such group we call it **non-split Cartan subgroup**.*

Lemma 4.3.6. *Let C be a Cartan subgroup and let N be its normaliser (i.e., the maximum group such that $C \subseteq N$ and C is normal in N). Then,*

$$N/C \simeq \{\pm 1\}.$$

Proof. We will prove it for C non-split, without loss of generality we can assume that C is the group of non-singular diagonal matrices D, by conjugation the result will follow. Let $M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in D$ and let $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in N$. Hence, $g^{-1}Mg$ must be diagonal, this gives that $ca = 0$ and $bd = 0$, $a = b = 0$ implies that $g \notin \mathrm{GL}_2(\mathbb{Z})$ and the same with $c = d = 0$. This implies that $a = d = 0$ or $b = c = 0$. Since $g \in \mathrm{GL}_2(\mathbb{Z})$ this implies that $a = d = \pm 1$ or $b = c = \pm 1$ so $g = \pm id$ or $g = \pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ so

$$N = D \cup \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} D$$

then it follows that $N/C \simeq \{\pm 1\}$.

If $C = \alpha^{-1}D\alpha$ where D is the group of diagonal matrices, hence

$$\begin{aligned} N &= \left\{g \in \mathrm{GL}_2(\mathbb{Z}) : \ g^{-1}Mg \in C, \ \forall M \in C\right\} = \left\{g \in \mathrm{GL}_2(\mathbb{Z}) : \ g^{-1}\alpha^{-1}E\alpha g \in \alpha^{-1}D\alpha, \ \forall E \in D\right\} \\ &= \left\{g \in \mathrm{GL}_2(\mathbb{Z}) : \ \alpha g^{-1}\alpha^{-1}E\alpha g\alpha^{-1} \in D, \ \forall E \in D\right\} = \alpha^{-1}N_D\alpha = \alpha^{-1}\left(D \cup \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} D\right)\alpha \\ &= \alpha^{-1}D\alpha \cup \alpha^{-1}\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\alpha\alpha^{-1}D\right)\alpha = C \cup \alpha^{-1}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\alpha C. \end{aligned}$$

From which the result follows. □

Lemma 4.3.7 (Cf. [SD73] Lemma 2.)**.** *Let ℓ be a prime number and G a subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$. If $\ell \mid |G|$ then G is contained in a Borel subgroup or $\mathrm{SL}_2(\mathbb{F}_\ell)$ is contained in G.*

Assume $\gcd(\ell, |G|) = 1$, and consider H de image of G in $\mathrm{PGL}_2(\mathbb{F}_\ell)$, then one of these three cases happens

1. *H is cyclic and G is contained in a Cartan subgroup.*
2. *H is dihedral and G is contained in the normaliser of a Cartan subgroup but not in the Cartan subgroup.*
3. *$H \simeq A_4$, $H \simeq S_4$ or $H \simeq A_5$.*

In the second case ℓ must be odd in the third ℓ must be prime to 6 if $H \simeq A_4$ or $H \simeq S_4$ and prime to 30 if $H \simeq A_5$.

Corollary 4.3.8. *Let*

$$\rho_\ell : \mathrm{Gal}(K_\ell/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\mathbb{Z}_\ell)$$

be a Galois representation such that $\det(\rho_\ell) = \chi^{k-1}$ for some integer k. Let G be the image of ρ composed with the reduction modulo ℓ, take H to be the image of G in $\mathrm{PGL}_2(\mathbb{F}_\ell)$. If $\mathrm{SL}_2(\mathbb{F}_\ell) \nsubseteq G$, then one of the following three situations happens

1. *G is contained in a Borel subgroup.*
2. *G is contained in the normaliser of a Cartan subgroup but not in the Cartan subgroup.*
3. *$H \simeq S_4$.*

Theorem 4.3.9. *Let $f \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ with integral Fourier coefficients and $a_1(f) = 1$. Assume that it is a Hecke form. Let*

$$\rho_\ell : \mathrm{Gal}(K_\ell/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\mathbb{Z}_\ell)$$

the Galois representation provided by Theorem 4.2.2. Assume that ℓ is an exceptional prime for f. Then, from the above corollary we have three cases:

1. *There exists m such that $a_n(f) \equiv n^m \sigma_{k-1-2m}(n) \pmod{\ell}$ for all n such that $\gcd(n, \ell) = 1$.*
2. *$a_n(f) \equiv 0 \pmod{\ell}$ if $\left(\frac{n}{\ell}\right) = -1$.*
3. *$p^{1-k} a_p^2 \equiv 0, 1, 2, 4 \pmod{\ell}$ for all primes p different from ℓ.*

Proof. 1. We can assume that our Borel subgroup involved is the one of the upper diagonal matrices. This means, for all $\sigma \in \mathrm{Gal}(K_\ell|\mathbb{Q})$

$$\overline{\rho}_\ell(\sigma) = \begin{pmatrix} \alpha(\sigma) & \beta(\sigma) \\ 0 & \gamma(\sigma) \end{pmatrix},$$

where $\alpha, \beta, \gamma : \mathrm{Gal}(K_\ell|\mathbb{Q}) \longrightarrow \mathbb{F}_\ell$ are continuous.

From theorem 4.2.2 we know $\alpha \cdot \gamma = \det(\overline{\rho}) = \overline{\chi}_\ell$. Hence, $\alpha = \overline{\chi}_\ell^m$ for some integer m, and $\gamma = \overline{\chi}_\ell^{k-1-m}$.

Taking $\sigma = \mathrm{Frob}(p)$, again by theorem 4.2.2 we have that the trace of the matrix of $\rho(\mathrm{Frob}(p))$ is $a_p(f)$ for any prime p different from ℓ, then

$$a_p(f) \equiv \alpha(\mathrm{Frob}(p)) + \gamma(\mathrm{Frob}(p)) \pmod{\ell}.$$

Using that $\det(\rho_\ell(\mathrm{Frob}(p))) = p^{k-1}$

$$\alpha(\mathrm{Frob}(p)) = p^m \quad \text{and} \quad \gamma(\mathrm{Frob}(p)) = p^{k-1-m}.$$

Therefore,

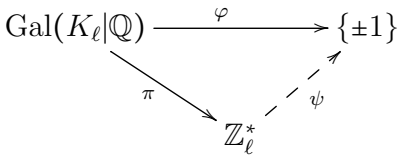
$$a_p(f) \equiv p^m + p^{k-1-m} \equiv p^m(1 + p^{k-1-2m}) = p^m \sigma_{k-1-2m}(p) \pmod{\ell},$$

and since f is a Hecke form, we deduce the result for any n coprime to ℓ. Note that the m does not depend on the choice of p, neither it depends on n.

2. Assume $\ell > 2$ (which is possible since all proper subgroups of $\mathrm{SL}_2(\mathbb{F}_2)$ are inside a Cartan group or a Borel group). Let C be the Cartan group provided by the above corollary and N its normaliser. Consider

$$\mathrm{Gal}(K_\ell|\mathbb{Q}) \longrightarrow N \longrightarrow N/C \simeq \{\pm 1\}$$

which is surjective by hypothesis. The image is abelian, then our morphism factors through $\mathrm{Gal}(K_\ell^{\mathrm{ab}}|\mathbb{Q}) \simeq \mathbb{Z}_\ell^*$:



but the only surjective and continuous morphism from $\mathbb{Z}_\ell^*$ to $\{\pm 1\}$ is the one whose kernels are the squares modulo ℓ. Then, $\varphi(\mathrm{Frob}(p)) = 1$ if and only if $\overline{\rho}_\ell(\mathrm{Frob}(p)) \in C$ and this happens if, and only if, p is a square in $\mathbb{F}_\ell$.

Let $\alpha \in N \setminus C$, doing a field extension if necessary α interchanges two one-dimensional subspaces of the space on which it operates and therefore it can be put as a

$$\alpha = \begin{pmatrix} 0 & * \\ * & 0 \end{pmatrix}$$

and has trace equal to 0. Hence, if p is quadratic non-residue modulo ℓ by 4.2.2

$$a_p(f) \equiv 0 \pmod{\ell}.$$

By the same argument as in 1 it follows the result for n.

3. If H is isomorphic to S_4 then all elements in H have order $1, 2, 3$ or 4, this is because the order of S_4 is 4! and it is not cyclic ($8|4!$ but an element of order 8 multiplied to an element of order 3 would be an element of order 4!). So the elements of H have characteristic roots equal to $\mu\lambda$, $\mu^{-1}\lambda$ with $\mu^n = 1$ for $n = 2, 4, 6$ or 8 and λ a number. Computing the characteristic polynomial in each case it is easy to find the result.

□

Remark 4.3.10. *Recall that this theorem talks about cusps forms with integral coefficients in the full modular group. Following the dimension formula for cusps forms (cf. [DS05] Theorem 3.5.2) we know that the only k for which* $\dim(\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))) = 1$ *are* $12, 16, 18, 20, 22$ *and* 26 *in which cases we can assure by Theorem 2.5.7 that there are cusps forms with the hypothesis in this last theorem. For $k \geq 28$ or $k = 24$ dimension is* 2 *or more and the Fourier coefficients are integers of degree at most the dimension. Which means we cannot assure the existence of cusps forms satisfying the hypothesis in this theorem.*

4.4 Exceptional primes

Given that we know now that there are congruences for modular forms modulo an exceptional prime, the natural question is about how is the set of exceptional primes, i.e., if they are infinite or not.

Lemma 4.4.1 (Cf.[SD73] Lemma 7.)**.** *Let f be a modular form for the full modular group with Fourier coefficients $\{a_n\} \subseteq \mathbb{Z}$. Assume $a_0 = 0$, $a_1 = 1$ and that it is a Hecke form. Let ℓ be a non exceptional prime for f, let $N \subseteq \mathbb{Z}_\ell$ and $N^* \subseteq \mathbb{Z}_\ell^*$ non empty subsets. The set*

$$\{p \text{ prime } : p \in N^* \text{ and } a_p \in N\}$$

has positive density.

Lemma 4.4.2. *With the same notation that in Theorem 4.3.9. We have that*

1. *Case 1 can only happen if $2m < \ell < k$ or $m = 0$ and ℓ divides the numerator of B_k.*
2. *Case 2 can only happen if $\ell < 2k$.*

Proof. Assume $\ell > 3$, which means case 1 is equivalent to

$$a_p \equiv p^m + p^{k-1-m} \pmod{\ell},$$

the exponents are only significant modulo $\ell - 1$ so we can reduce them to their class modulo $\ell - 1$ and interchange them if necessary to find

$$a_p \equiv p^m + p^{m'} \pmod{\ell}$$

such that $0 \leq m < m' < \ell - 1$, with $m + m' \equiv k - 1 \pmod{\ell - 1}$, $m \neq m'$ since their sum is odd. So with an argument similar to that in 4.3.9 we get

$$a_n \equiv n^m \sigma_{m-m'}(n) \pmod{\ell}$$

for all n coprime to ℓ.

This translates into

$$\theta \overline{f} = \theta^{m+1} \overline{G}_{m'-m+1},$$

putting an extra θ in both sides annihilate the terms where $n \equiv 0 \pmod{\ell}$ and the other cases rest the same since n coprime to ℓ means they are invertible modulo ℓ. This can be done if $m \neq 0$ and $m' \neq \ell - 2$, in which case the constant term of $G_{m'-m+1}$ is not in $\mathfrak{o}$, so we have instead $pa_p = 1 + p \pmod{\ell}$ and then $na_n \equiv \sigma_1(n) \pmod{\ell}$ if $(n, \ell) = 1$. So

$$\theta \overline{f} = \theta^{\ell-1} \overline{G}_2 = \theta^{\ell-1} \overline{G}_{\ell+1}. \tag{4.1}$$

By lemma 1.6.10 $\omega(\theta\overline{f}) \leq k+\ell+1$, and $\omega(\overline{G}_k) = k$ for even k and $2 < k < \ell - 1$, so iterating on

$$\theta\overline{f} = \theta^{m+1}\overline{G}_{m'-m+1}, \tag{4.2}$$

we are always in the case of equality. Then, if $m' - m > 1$ the filtration of the right hand is exactly $(m'-m+1)+(m+1)(\ell+1)$, so using $\omega(\theta\overline{f}) \leq k+\ell+1$, we get to

$$(m'-m+1)+(m+1)(\ell+1) \leq k+\ell+1.$$

So, operating

$$\begin{aligned} m'-m+1+m\ell+m+\ell+1 &\leq k+\ell+1 \\ m'+m\ell+\ell+2 &\leq k+\ell+1 \\ m'+m\ell+1 &\leq k \end{aligned}$$

given $m - m' > 1$ and $\ell - 2 > m - m'$.

If $\ell > k$, then by what we have seen above, $m + m' \geq k - 1$, so using our inequality

$$\begin{aligned} m'+m\ell &\leq k-1 \\ m'+m-m+m\ell &\leq k-1 \\ (m'+m)+m(\ell-1) &\leq k-1 \\ (k-1)+m(k-1) &\leq k-1 \\ (k-1)(1+m) &\leq k-1 \end{aligned}$$

which is only possible if $m = 0$ and then $m' = k-1$ and $\omega(\overline{f}) = k$. Which makes equation (4.1) to

$$\theta(\overline{f}) = \theta(\overline{G}_k)$$

so $\theta(\overline{f}-\overline{G}_k) = 0$ and since $\omega(\overline{f}) = \omega(\overline{G}_k) = k$ we have that either $\omega(\overline{f}-\overline{G}_k) = k$ or either $\omega(\overline{f}-\overline{G}_k) = 0$. From lemma 1.6.10 we know

$$0 = \omega(0) = \omega(\theta(\overline{f}-\overline{G}_k)) \leq \omega(\overline{f}-\overline{G}_k)+\ell+1$$

and we have inequality if, and only if $\omega(\overline{f}-\overline{G}_k) \equiv 0 \pmod{\ell}$, and since $\omega(\overline{f}-\overline{G}_k)$ is either k which is even or 0 we deduce that it is 0. Which means $\overline{f}-\overline{G}_k = 0$ because $\mathfrak{M}_0 = \{0\}$. Since f was a cusp, $a_0(f) = 0$, and $a_0(G_k) = \frac{B_k}{2_k}$ so from $f \equiv G_k \pmod{\ell}$ we deduce that ℓ divides the numerator of B_k.

Similarly, for equation 4.2 when $m'-m = 1$ we can calculate the filtration. The right side is $\overline{G}_2$ and $\overline{B}(\overline{E}_4,\overline{E}_6) = \overline{E}_2$ so, since there are no modular forms of weight 2 and $\omega(\overline{G}_2) \neq 0$

$$\begin{aligned} \omega(\theta^{m+1}\overline{G}_2) =& \omega(\theta^m\overline{G}_2)+\ell+1 = \omega(\theta^{m-1}\overline{G}_2)+2\ell+2 = \cdots = \omega(\overline{G}_2)+(m+1)(\ell+1) \\ =& \ell+1+(m+1)(\ell+1) = (m+2)(\ell+1). \end{aligned}$$

Now, for 4.1

$$\begin{aligned}\omega(\theta^{\ell-1}\overline{G}_{\ell+1}) = \omega(\theta^{\ell-2}\overline{G}_{\ell+1}) + \ell + 1 = \cdots = \omega(\overline{G}_{\ell+1}) + (\ell-1)(\ell+1)\\ = \ell + 1 + (\ell-1)(\ell+1) = \ell(\ell+1).\end{aligned}$$

Comparing with $\theta\overline{f}$ we get

$$(m+2)(\ell+1) = \omega(\theta\overline{f}) \leq \omega(\overline{f}) + \ell + 1 = k + \ell + 1$$

and hence

$$(m+1)(\ell+1) \leq k \quad \text{if } m' - m = 1.$$

For 4.1

$$\ell(\ell+1) = \omega(\theta\overline{f}) \leq \omega(\overline{f}) + \ell + 1 = k + \ell + 1,$$

then

$$\ell^2 - 1 \leq k \quad \text{if } m = 0,\ m' = \ell - 2.$$

Both of them imply $\ell < k$.

Consider now ℓ to be a prime of the second type. As we have done with the first case, we can write it using θ

$$\theta\overline{f} = \theta^{(\ell+1)/2}\overline{f},$$

assume $\ell > 2k$, hence $\omega(\overline{f}) = k$, so the filtration of the left hand side is exactly $k + \ell + 1$ but the right hand side is $k + \frac{(\ell+1)^2}{2}$, which implies

$$k + \ell + 1 = k + \frac{(\ell+1)^2}{2} \implies \ell = 1,$$

which is a contradiction.

This proves the lemma since ℓ is odd and k is even, so $\ell = k$ and $\ell = 2k$ are impossible situations.

□

Theorem 4.4.3. *With the same notations, there are finitely many exceptional primes for f. Those of cases 1 and 2 can be explicitly determined and it can be explicitly determined a finite set containing the primes in case 3.*

Proof. The above lemma shows that the exceptional primes for cases 1 and 2 are finite. Consider then the third case.

The idea of Swinnerton-Dyer is to produce a finite list containing all such primes but the list will sometimes contain some primes non-exceptional.

Consider $p \neq 2$ such that $a_p(f) \neq 0$, then if ℓ is an exceptional prime of those of type 3, ℓ divides one of

$$a_p^2,\ a_p^2 - p^{k-1},\ a_p^2 - 2p^{k-1},\ a_p^2 - 4p^{k-1},$$

or $\ell = p$. Since k is even at least one of them is non zero, so this gives a finite list of such ℓ. □

In [SD73], Swinnerton-Dyer gives a list of exceptional primes for the first forms.

For example the exceptional primes for Δ of type 1 such that $\ell < k$ are $2, 3, 5, 7$ for which $m = 0, 0, 1, 1$ respectively. And the only exceptional prime of type 1 with $\ell > k$ (i.e. dividing the numerator of B_{12}) is, as expected 691. The only exceptional prime of type 2 for Δ is $\ell = 23$.

For primes of the third type there is $\ell = 59$ for $E_4\Delta$. None of the other cusps have no such a primes. For other examples cf. [SD73] Corollary to Theorem 4, pp. 31-32.

It is remarkable that the results for primes 2 and 3 are not a consequence of what we have prove since we were assuming all the proof long that $\ell > 3$, but follow from

$$\tau(p) \equiv 0 \pmod 2, \quad \tau(p) \equiv p + p^2 \pmod 3,$$

and Proposition 1.6.11.

4.5 Congruences for modular forms on $\Gamma_0(N)$

If in the last two sections we have presented the cases for level $N = 1$ now we present the approach that K. Ono followed in [Ono94] using the same strategies that Swinnerton-Dyer developed. We should realise that all information about the congruences between modular forms come from the knowledge of the image of certain Galois representations linked to our cusps forms whose existence was conjectured by Serre and proved by Deligne in [Del69].

On the same way,

Theorem 4.5.1 (Cf. [Del69]). *Let $f \in \mathcal{S}_k(N, \varepsilon)$ a normalised Hecke form with Fourier coefficients $a_1 = 1$ and $\{a_n\}_n \subseteq \mathcal{O}$. Then,for any prime ℓ there is a continuous homomorphism*

$$\rho_f : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\mathbb{Z}_\ell),$$

unramified outside ℓN such that $\rho_f(\mathrm{Frob}(p))$ has characteristic polynomial

$$X^2 - a_p X + \varepsilon(p)p^{k-1},$$

for each prime p not dividing ℓN.

Theorem 4.5.2. *Assume $\ell = N$, then ρ_ℓ factors through K (where as in the two sections before, it is the maximal field extension of $\mathbb{Q}$ only ramified at ℓ). In this case, $\overline{\rho_\ell}(\mathrm{Gal}(K|\mathbb{Q}))$ cannot be contained in a non-split Cartan subgroup without being contained in a Borel subgroup C.*

Proof. Assume that the image $\overline{\rho_\ell}(\mathrm{Gal}(K|\mathbb{Q}))$ is contained in a non-split Cartan subgroup C. Since C is abelian, we have that

$$\begin{array}{ccc} \mathrm{Gal}(K|\mathbb{Q}) & \xrightarrow{\overline{\rho}_\ell} & C \\ {\scriptstyle\pi}\downarrow & \nearrow & \\ \mathrm{Gal}(K^{\mathrm{ab}}|\mathbb{Q}) & & \end{array}$$

is a commutative diagram. Since all factor groups of $\mathrm{Gal}(K^{\mathrm{ab}}|\mathbb{Q})$ have order dividing $\ell^n(\ell-1)$ for some n, it follows that $|\overline{\rho_\ell}(\mathrm{Gal}(K|\mathbb{Q}))|$ divides $\ell-1$, so the matrices in $\overline{\rho_\ell}(\mathrm{Gal}(K|\mathbb{Q}))$ have eigenvalues in $\mathbb{F}_\ell$ since their characteristic polynomials divide $X^{\ell-1}-1$. Due that they commute, they can all be diagonalised at the same time so $\mathrm{Gal}(K_\ell|\mathbb{Q})$ is contained in a Borel subgroup. □

Corollary 4.5.3 (Cf. [Ono94] Corollary 3.1.)**.** *Let $f \in \mathcal{S}_k(N,\varepsilon)$ be a normalised Hecke form and let $G = \overline{\rho_\ell}(\mathrm{Gal}(\overline{\mathbb{Q}}|\mathbb{Q})) \subseteq \mathrm{GL}_2(\mathbb{F}_\ell)$, and H the image of G in $\mathrm{PGL}_2(\mathbb{F}_\ell)$. Then, either*

1. *G is contained in a Borel subgroup.*
2. *G is contained in the normaliser of a Cartan subgroup but not in the Cartan subgroup itself.*
3. *$H \simeq S_4$.*
4. *$H \simeq A_4$.*
5. *$H \simeq A_5$.*

Case 2 can happen if $\ell > 2$, case 3 and 4 if $\ell > 3$ and case 5 if $\ell > 5$.

Remark 4.5.4. *Recall that the normaliser of a subgroup $H \subseteq G$ is the maximum N such that $H \subseteq N \subseteq G$ with $H \lhd N$. So,*

$$N = \{g \in G : \ gH = Hg\}.$$

Theorem 4.5.5. *Let $f \in \mathcal{S}_k(N,\varepsilon)$ be a normalised Hecke form, let $G = \overline{\rho_\ell}(\mathrm{Gal}(\overline{\mathbb{Q}}|\mathbb{Q})) \subseteq \mathrm{GL}_2(\mathbb{F}_\ell)$ be its attached Galois representation and H the image of G in $\mathrm{PGL}_2(\mathbb{F}_\ell)$. Then,*

1. *If G is contained in a Borel subgroup and ε is trivial, there is an integer m such that it $(\ell, m) = 1$*
$$a_n \equiv n^m \sigma_{k-1-2m} \pmod{\ell}.$$

2. *If G is contained in the normaliser of a Cartan subgroup but not in the Cartan subgroup itself, if $(n, \ell) = 1$ and n is a quadratic non-residue modulo ℓ then*
$$a_n \equiv 0 \pmod{\ell}.$$

3. *If $H \simeq S_4$, then if $p \nmid \ell$*
$$\varepsilon^{-1}(p)p^{1-k}a_p^2 \equiv 0, 1, 2, 4 \pmod{\ell}.$$

4. *If $H \simeq A_4$, then if $p \nmid \ell$*
$$\varepsilon^{-1}(p)p^{1-k}a_p^2 \equiv 0, 1, 4 \pmod{\ell}.$$

5. *If $H \simeq A_5$, then if $p \nmid \ell$*
$$\left(\varepsilon^{-1}(p)p^{1-k}a_p^2 - \frac{3}{2}\right)^2 \equiv \frac{1}{4}, \frac{5}{4}, \frac{9}{4}, \frac{25}{4} \pmod{\ell}.$$

Case 2 can happen if $\ell > 2$, case 3 and 4 can occur if $\ell > 3$ and case 5 if $\ell > 5$.

Proof. The proof of the first three statements is very similar to that of Theorem 4.3.9 and statement 4 is very similar to case 3, hence we do not do it.

Consider case 5. Let $A \in \mathrm{PGL}_2(\mathbb{F}_\ell)$ a order 5 matrix, then there is $A'' \in \mathrm{GL}_2(\mathbb{F}_\ell)$ a representative of A. Then, $\det(A''^5) = k^2$, $k \in \mathbb{F}_\ell$. So there is $A' \in \mathrm{GL}_2(\mathbb{F}_\ell)$ of the same class of A'' with determinant 1. And hence, its characteristic polynomial divides the cyclotomical polynomial of order 5, that in $\mathbb{F}_{\ell^2}$ factors
$$T^4 + T^3 + T^2 + T + 1 = (T^2 + aT + 1)(T^2 + bT + 1)$$
with $a = (1 + \sqrt{5})/2$ and $b = (1 - \sqrt{5})/2$. Relating the trace and the determinant the formula follows.

□

Chapter 5

Congruences for special values of L-functions

We will study a very important result by H. Hida [Hid81a] using L functions which relates some special values of the associated L-functions and their prime divisors p with the existence of non-conjugate to f under $G_{\mathbb{Q}}$ Hecke form congruent to f modulo some prime above p.

In Hida's paper he uses Parabolic Cohomology introduced by Shimura in [Shi73] which we shall not comment since it exceed the goal and scope of this thesis.

5.1 L-functions

In the study of congruence primes, Hida (in [Hid81a], [Hid81b]) showed that thay can be studied using the adjoint L-function. In order to understand the statement of the results due to Hida we begin with L-functions.

First, we can start with the classical definition of an L-function associated to a Dirichlet character modulo m

Definition 5.1.1. *Let m be a positive integer and χ a Dirichlet character modulo m. We define the L function associated to χ to be*

$$L(s,\chi) = \sum_{n\geq 1} \frac{\chi(n)}{n^s}.$$

Lemma 5.1.2. *The L-function associated to χ is absolutely convergent for $\Re(s) > 1$*

Proof. $|L(s,\chi)| \leq \sum_{n\geq 1} \left|\frac{\chi(n)}{n^s}\right| \leq \sum_{n\geq 1} \frac{1}{|n^s|} = \sum_{n\geq 1} n^{-\Re(s)}$ which converges if $\Re(s) > 1$. □

In a similar way, we can do the following construction.

Definition 5.1.3. *Let k, N be integers and $f \in \mathcal{M}_k(\Gamma_1(N))$. Let $f = \sum_{n\geq 0} a_n q^n$ be its Fourier expansion around infinity. We define its L-function to be the complex variable function*

$$L(s, f) = \sum_{n\geq 1} a_n n^{-s}.$$

Proposition 5.1.4 (Cf. [DS05] Proposition 5.9.1.)**.** *Let k, N be integers and $f \in \mathcal{M}_k(\Gamma_1(N))$. If f is a cusp, then $L(s, f)$ converges absolutely for any s such that $\Re(s) > \frac{k}{2}+1$ and, if f is not a cusp then $L(s, f)$ converges absolutely for any s satisfying $\Re(s) > k$.*

Proposition 5.1.5. *Let N, k be integers and χ a Dirichlet character modulo N. If $f \in \mathcal{M}_k(N, \chi)$ with $f = \sum_{n\geq 0} a_n q^n$ its Fourier expansion around infinity. The following are equivalent:*

- *f is a normalised Hecke form.*
- *$L(s, f)$ has an Euler product*

$$L(s, f) = \prod_{p \text{ prime}} (1 - a_p p^{-s} + \chi(p)p^{k-1-2s})^{-1}.$$

Proof. We know by Proposition 2.4.3 that the Fourier coefficients a_n must satisfy

$$\begin{cases} a_1 = 1, \\ a_{p^r} = a_p a_{p^{r-1}} - \chi(p)p^{k-1} a_{p^{r_2}}, \text{ for all prime } p \text{ and } r \geq 2, \\ a_{mn} = a_n a_m, \text{ when } n, m \text{ are coprime.} \end{cases}$$

We will prove that the second condition in the theorem is equivalent to these three. Fix a prime p, multiply the second condition by p^{-rs} and sum over $r \geq 2$

$$\sum_{r\geq 0} a_{p^r} p^{-rs}(1 - a_p p^{-s} + \chi(p)p^{k-1-2s}) = a_1 + (1 - a_1)a_p p^{-s},$$

by the first condition of the Hecke forms we get

$$\sum_{r\geq 0} a_{p^r} p^{-rs}(1 - a_p^{-s} + \chi(p)p^{k-1-2s}) = 1.$$

Conversely, assume this last equality holds and let $s \to \infty$, then $a_1 = 1$, so the last two equations hold and $a_1 = 1$, this implies condition 2 of the Hecke forms. Summarising, conditions 1 and 2 of the Hecke form characterisation are equivalent to

$$\sum_{r\geq 0} a_{p^r} p^{-rs} = (1 - a_p^{-s} + \chi(p)p^{k-1-2s})^{-1}, \quad \text{for prime } p.$$

Write now $p^r \| n$ if $p^r | n$ but $p^{r+1} \nmid n$, then, if g is a function of prime powers

$$\prod_p \sum_{r\geq 0} g(p^r) = \sum_{n\geq 1} \prod_{p^r \| n} g(p^r).$$

So if the last equality holds by this remark and with the third condition of the characterisation of Hecke forms we can compute

$$L(s,f) = \sum_{n\geq 1} a_n n^{-s} = \sum_{n\geq 1} \left(\prod_{p^r \| n} a_{p^r} \right) n^{-s},$$

because of the third condition. Now, by our observation

$$L(s,f) = \sum_{n\geq 1} \prod_{p^r \| n} a_{p^r} p^{-rs} = \prod_p \sum_{r\geq 0} a_{p^r} p^{-rs} = \prod_p (1 - a_p p^{-s} + \chi(p) p^{k-1-2s})^{-1}$$

as we wanted to see.

Conversely, given the Euler product expansion using the geometric series formula

$$L(s,f) = \prod_p (1 - a_p p^{-s} + \chi(p) p^{k-1-2s})^{-1} = \prod_p \sum_{r\geq 0} b_{p,r} p^{-rs}$$

for some $b_{p,r}$. And using the observation again

$$L(s,f) = \sum_{n\geq 1} \prod_{p^r \| n} b_{p,r} p^{-rs} = \sum_{n\geq 1} \left(\prod_{p^r \| n} b_{p,r} \right) n^{-s}.$$

This gives

$$a_n = \prod_{p^r \| n} b_{p,r}$$

which implies the third condition of the characterisation of Hecke forms. This implies too the equality above, and hence the first and second conditions. □

Example 5.1.6. *Let*

$$L(s,\Delta) = \sum_{n\geq 1} \frac{\tau(n)}{n^s} = \prod_{p \ prime} \frac{1}{(1 - \tau(p) p^{-s} + p^{11-2s})}.$$

5.2 Discriminants of Quadratic forms

Definition 5.2.1. *Let K be a field, we say that is a CM-field if there is $k|\mathbb{Q}$ totally real such that $K|k$ is totally imaginary and $[K : k] = 2$.*

Definition 5.2.2. *Let K be a CM-field or a totally real field of finite degree over $\mathbb{Q}$. Let V a finite dimensional K-vector space and T a skew-symmetric non-degenerate $\mathbb{Q}$-bilinear form which takes values on $\mathbb{Q}$ such that for any $a \in K,\ x, y \in V$*

$$T(ax, y) = T(x, \overline{a}y),$$

where $\overline{a}$ is the complex conjugate of a.

Let Λ be a lattice of V over $\mathbb{Z}$ and $\{e_1, \dots, e_m\}$ be a basis. Take R the matrix of T for this basis, we define the discriminant of T with respect to Λ.

$$d(T, \Lambda) = \det(R).$$

Notice that this definition does not depend on the basis of Λ we take since a change of basis can be expressed as an element in $\mathrm{SL}_m(\mathbb{Z})$ and hence, the determinant does not change. Moreover, the change of Λ to Λ' depends on a matrix M and M^t its traspose. So $d(T) = d(T, \Lambda)$ in $\mathbb{Q}^*/\mathbb{Q}^{*2}$ does not depend on Λ.

Definition 5.2.3. *Let Λ, Λ' two lattices as above, then define*

$$[\Lambda : \Lambda'] = \frac{[\Lambda : \Lambda \cap \Lambda']}{[\Lambda' : \Lambda \cap \Lambda']} = [\Lambda : \Lambda'].$$

This is an abuse of notation because the left hand side represents the index defined above and the right hand side represents the usual concept of index of groups.

Remark 5.2.4. *The former definition may seem artificial but notice that if $\Lambda' \subset \Lambda$,*

$$[\Lambda : \Lambda'] = \frac{[\Lambda : \Lambda \cap \Lambda']}{[\Lambda' : \Lambda \cap \Lambda']} = \frac{[\Lambda : \Lambda']}{[\Lambda' : \Lambda']}$$

which is the usual definition.

Proposition 5.2.5. *Let Λ be a lattice in V and let Λ^* be the dual, then*

$$|d(T, \Lambda)| = [\Lambda^* : \Lambda].$$

The proof is just linear algebra.

5.3 Petersson Inner Product

Theorem 5.3.1 (Cf.[Shi75] Theorem 1.)**.** *If f is a newform form of weight k and nebentype χ, if we define α_p, β_p such that*

$$\alpha_p + \beta_p = a_p, \quad \alpha_p\beta_p = \begin{cases} \chi(p)p^{k-1}, & p \nmid N, \\ 0, & p|N. \end{cases} \quad , \quad p \text{ prime}.$$

Then, if ρ is a Dirichlet character modulo M we define

$$L(s,f,\rho)=\prod_p\left(\left(1-\rho(p)\alpha_p^2p^{-s}\right)\left(1-\rho(p)\alpha_p\beta_pp^{-s}\right)\left(1-\rho(p)\beta_p^2p^{-s}\right)\right)^{-1}.$$

$L(s,f,\rho)$ converges for $\Re(s)>>0$ and can be extended to a meromorphic function. It only can have simple poles in $s=k,k-1$. Indeed, it has a pole in $s=k$ if, and only if

- *$\chi\rho$ is a non-trivial Dirichlet character of order 2.*
- *If $g=\sum_{n\geq 0}\overline{\rho(n)}\overline{a_n}q^n$ and $\langle f,g\rangle\neq 0$, (it can be seen that $g\in\mathcal{S}_k\Gamma_1(M^2N)$.)*

Proposition 5.3.2. *Let $f=\sum_{n\geq 1}a_nq^n\in\mathcal{S}_k(\Gamma_1(N))$, $g=\sum_{n\geq 1}b_nq^n\in\mathcal{S}_l(\Gamma_1(N))$ be two normalised Hecke forms. Define α_p,β_p for f as in the theorem above and α_p',β_p' for g. Then,*

$$D(s,f,g):=\sum_{n\geq 1}\frac{a_nb_n}{n^s}=\prod_p\left(\left(1-\rho(p)\alpha_p\alpha_p'p^{-s}\right)\left(1-\rho(p)\alpha_p\beta_p\alpha_p'\beta_p'p^{-2s}\right)\left(1-\rho(p)\beta_p\beta_p'p^{-s}\right)\right.$$
$$\left.\left(1-\rho(p)\alpha_p\beta_p'p^{-s}\right)\left(1-\rho(p)\beta_p\alpha_p'p^{-s}\right)\right)^{-1}.$$

Proposition 5.3.3. *With the above notation,*

$$\frac{\pi^2}{6}\left(\prod_{p|N}(1-p^{-2})\right)\mathrm{Res}_{s=k}D(s,f,g)=\left(\prod_{p|N}(1-p^{-1})\right)L(k,f,\overline{\chi}).$$

Proof. In [Shi75] Shimura proves that if $\zeta_N(s)=\sum_{n\geq 1,(n,N)=1}\frac{1}{n^s}$, then

$$\zeta_N(2s-2k+2)D(s,f,g)=\zeta_N(s-k+1)L(s,f,\overline{\chi}),$$

where $g=\sum_{n\geq 1}\overline{\chi(n)}\overline{a_n}q^n\in\mathcal{S}_k(N^2)$. Then, if we compare the residues of the poles at $s=k$, noting that $L(k,f,\overline{\chi})$ is holomorphic at $s=k$, we have that

$$\zeta_N(2)\mathrm{Res}_{s=k}D(s,f,g)=L(k,f,\overline{\chi})\mathrm{Res}_{s=1}\zeta_N(s).$$

We know that $\zeta(2)=\frac{\pi^2}{6}$ and $\mathrm{Res}_{s=1}\zeta(s)=1$. Then, substituting we get the result. □

Proposition 5.3.4. *Let N be the conductor of f a newform form and M the conductor of χ*

$$D(s,f,f_p)=\left(\prod_{p\in A}(1-p^{k-1-s})^{-1}\right)D(s,f,g),$$

where A is the set of all primes dividing N such that $N_p=M_p$, where N_p,M_p are the maximum powers of p dividing N and M respectively.

Proof. Keep the notation of the proposition above. We know that $\overline{a_n} = \overline{\chi(n)}a_n$ for any n coprime to N, hence if $f_p(z) = \sum_{n\geq 1} \overline{a_n} e^{nz}$, we have that

$$D_N(s, f, f_p) = D_N(s, f, g),$$

where the subscript means that we avoid the N-factors. Then, for any $p|N$

$$\begin{cases} a_p\overline{a_p} = p^{k-1}, & \text{if } p|N \text{ and } N_p = M_p, \\ a_p\overline{a_p} = p^{k-2}, & \text{if } p|N \text{ and } N_p = p,\ M_p = 1, \\ a_p = 0, & \text{if } p^2|N \text{ and } N_p \neq M_p. \end{cases}$$

The proof of this fact can be seen in [KD76].

Since we have that f is a newform, this implies

$$\overline{a_n} = \overline{\chi(n)} = a_n, \quad \text{for } n \text{ coprime to } M.$$

Therefore, when a prime p is coprime to M the Euler p-factor of $D(s, f, f_p)$ coincides with the p-factor of $D(s, f, g)$ defined above. If $M_p \neq 1$ and $N_p \neq M_p$ the Euler p-factor of both are 1 since $a_p = 0$. Thus we deduce that

$$D(s, f, f_p) = \left(\prod_{p\in A}(1 - p^{k-1-s})^{-1}\right) D(s, f, g),$$

as we wanted to see. □

Proposition 5.3.5 (Cf. [Shi76] (2.5))**.**

$$\mathrm{Res}_{s=k} D(s, f, f_p) = \frac{6(4\pi)^k}{\pi(k-1)!\delta(N)\prod_{p|N}(1 - p^{-2})}\langle f, f\rangle.$$

Theorem 5.3.6. *Let f be a newform of weight k, conductor N and nebentype χ, let M be the conductor of χ. Then,*

$$L(k, f, \overline{\chi}) = \frac{2^{2k}\pi^{k+1}}{(k-1)!\delta(N)NM\varphi(N/M)}\langle f, f\rangle.$$

Where φ denotes the Euler's totient function and $\delta(N) = 2$ if $N \leq 1$ and $\delta(N) = 1$ otherwise.

Proof. Using the above propositions

$$\frac{(4\pi)^k}{(k-1)!\left(\frac{\pi}{6}N^2\delta(N)\prod_{p|M}(1 - p^{-2})\right)}\langle f, f\rangle = \mathrm{Res}_{s=k} D(s, f, f_\rho) = \mathrm{Res}_{s=k}\frac{D(s, f, g)}{\prod_{p\in A}(1 - p^{k-1-s})}$$
$$= \frac{1}{\prod_{p\in A}(1 - p^{k-1-k})}\frac{\prod_{p|N}(1 - p^{-1})L(k, f, \overline{\chi})}{\frac{\pi^2}{6}\prod_{p|N}(1 - p^{-2})}.$$

Hence,

$$L(k,f,\overline{\chi}) = \frac{2^{2k}\pi^{k+1}\prod_{p\in A}(1-p^{-1})}{(k-1)!N^2\delta(N)\prod_{p|N}(1-p^{-1})}\langle f,f\rangle = \frac{2^{2k}\pi^{k+1}}{(k-1)!\delta(N)NM\varphi(N/M)}\langle f,f\rangle.$$

□

5.4 Discriminants and newforms

Let f be a newform of $\mathcal{S}_k(\Gamma_1(N))$ of nebentype χ. We denote by $K|\mathbb{Q}$ the extension which contains the Fourier coefficients of f at $i\infty$, let r denote the degree of this extension.

Theorem 5.4.1 (Cf. [Hid81a] 3.3)**.** *Assume $k \geq 2$ and let Λ^* be the dual of Λ in $W(\mathbb{Q})$ under T. Then*

$$d(f) = [\Lambda^* : \Lambda] = \left(\frac{2^{(k-2)r}\prod_\sigma \langle f^\sigma, f^\sigma\rangle}{u_f}\right)^2.$$

First of all, take G be the group of automorphisms of K and define the space

$$S(f) = \langle f^\sigma : \ \sigma \in G\rangle_{\mathbb{C}} \subseteq \mathcal{S}_k(\Gamma_1(N)).$$

Let's define now the first parabolic cohomology group. Let H be a group and M be a H-module. A 1**-cocycle** is a map

$$u : H \longrightarrow M,$$

such that for any $g, h \in H$, we have

$$u(gh) = u(g) + gu(h).$$

Let $Z(H,M)$ be the set of 1-cocycles of H with values in M and for any subset $P \subseteq H$,

$$Z(H,M)_P = \{u \in Z(H,M) : \ u(p) \in (p-1)M \quad \forall p \in P\}.$$

Let also $B(H,M) = \{u \in Z(H,M) : \ u(\alpha) = (\alpha-1)x, \quad x \in M, \alpha \in H\}.$

Then,

$$H^1_P(H,M) = Z_P(H,M)/B(H,M).$$

Lemma 5.4.2 (Cf. [Hid81a] 3.4)**.** *If Γ is a congruence subgroup and $\overline{\Gamma} = \Gamma/\Gamma \cap \mathbb{Q}^*$, then there is an isomorphism*

$$\varphi : S_k(\Gamma) \longrightarrow H^1_P(\overline{\Gamma}, \mathbb{Z}^{n+1} \otimes_{\mathbb{Z}} \mathbb{R}).$$

From this isomorphism we induce a scalar product in $H^1_P(\overline{\Gamma}, \mathbb{Z}^{n+1} \otimes_{\mathbb{Z}} \mathbb{R})$ by

$$A_\Gamma(f,g) = (2i)^{n-1}\left(\langle f,g\rangle_\Gamma + (-1)^{n+1}\langle g,f\rangle\right),$$

and

$$\langle x,y\rangle_N = A_\Gamma(\varphi^{-1}(x), \varphi^{-1}(y)).$$

For now on, it is better to simplify notation and we will write $V(N:R)$ instead of $H^1_P(\overline{\Gamma}, \mathbb{Z}^{n+1} \otimes_{\mathbb{Z}} \mathbb{R})$. Consider $W_f(\mathbb{R})$ the image under ϕ of $S(f)$ inside $V(N:\mathbb{R})$. Define

$$L_f = W_f(\mathbb{R}) \cap V(N:\mathbb{Z}).$$

It is possible to see that L_f is a lattice in $W(\mathbb{R})$ which has a structure of vector space over K and

$$\langle L_f, L_f\rangle_N \subseteq e(N)^{-1}\mathbb{Z},$$

with $e(N) = 6, 2, 1$ if $N = 1, 2$ or ≥ 3 respectively. Let Y be the orthogonal complement of $W_f(\mathbb{R})$ in $V(N:\mathbb{R})$ (i.e., $V(N:R) = W_f(\mathbb{R}) \oplus Y$) under $\langle , \rangle_N$, $L_Y = Y \cap V(N:\mathbb{R})$ is a lattice.

$W_f(\mathbb{R})$ is stable under T_n and Y under T_n^*, and if $W_f(\mathbb{R})$ is stable under T_n^* Y is stable under the action of T_n, which happens if f is a newform. Consider then M_f, M_Y the projections of $V(N:\mathbb{Z})$, these are lattices and

$$L_f \subseteq M_f, \quad L_Y \subseteq M_Y$$

stable under the Hecke action and the dual Hecke action. Let T_f be the bilinear form induced by $\langle , \rangle_N$ in $W_f(\mathbb{R})$ and let $d(f) = d(T_f; L_f)$ defined in the second section. It is easy to see that $T_f(ax, y) = T_f(x, \overline{a}y)$ where $\overline{a}$ denotes complex conjugation. And, by the definition, T_f is symmetric if k is odd and skew-symmetric otherwise.

Let $B = \{f, f_2, \ldots, f_r\}$ be the conjugates of f under the action of $G_{\mathbb{Q}}$ where $r = [K : \mathbb{Q}]$. Then a basis of $S(f)$ over $\mathbb{R}$ is $B \cup iB$, using the above isomorphism define

$$\omega_j = \varphi(f_j), \quad \omega_{j+r} = \varphi(if_j), \quad 1 \leq j \leq r.$$

Which is a basis of $W_f(\mathbb{R})$ over $\mathbb{R}$. Let $\{\delta_1, \ldots \delta_{2r}\}$ be a basis over $\mathbb{Z}$ of L_f and let U be the matrix changing from this basis to the ω_j basis. Set

$$u(f) = \det(U),$$

u is well defined and does not depend on the choice of the basis of L_f.

Proof of the Theorem. Let $S = (\langle \omega_i, \omega_j \rangle_N)_{1 \le i.j \le 2r}$, then

$$\langle \omega_k, \omega_j \rangle_N = \begin{cases} A_\Gamma(f_k, f_j), & \text{if } 1 \le k, j \le r, \\ A_\Gamma(f_k, if_j), & \text{if } 1 \le k \le r,\ r+1 \le j \le 2r, \\ A_\Gamma(if_k, f_j), & \text{if } 1 \le j \le r,\ r+1 \le k \le 2r, \\ A_\Gamma(if_k, if_j), & \text{if } r+1 \le k, j \le 2r, \end{cases}$$

Since the conjugates are all different, we can find n such that $a_n(f_k) \neq a_n(f_j)$ and $(n, N) = 1$, since f is a newform this implies that $\langle f_k, f_j \rangle = 0$ if $i \neq j$, cf. [Miy71]. In particular we have

$$A_\Gamma(f_k, f_j) = A_\Gamma(if_k, if_j) = 0, \quad \text{if } i \neq j.$$

$$A_\Gamma(f_j, f_j) = A_\Gamma(if_j, if_j) = \begin{cases} (-1)^{(n-1)/2} 2^n \langle f_j, f_j \rangle_\Gamma, & \text{if } n \text{ is odd}, \\ 0, & \text{otherwise.} \end{cases}$$

$$A_\Gamma(f_j, if_j) = (-1)^{n+1} A_\Gamma(if_j, f_j) = \begin{cases} (-1)^{(n-2)/2} 2^n \langle f_j, f_j \rangle_\Gamma, & \text{if } n \text{ is even}, \\ 0, & \text{otherwise.} \end{cases}$$

Summarising, $\det(S) = 2^{2nr} \prod_\sigma \langle f^\sigma, f^\sigma \rangle_\Gamma^2$. Moreover, $S = UTU^t$, so

$$\det(S) = u^2 d(T_f; L_f) = 2^{2nr} \prod_\sigma \langle f^\sigma, f^\sigma \rangle_\Gamma^2.$$

And applying 5.2.5 the result follows. □

Definition 5.4.3. *Define*

$$c(f) = \frac{2^{(k-2)r} e(N) \prod_\sigma \langle f^\sigma, f^\sigma \rangle}{u(f)},$$

where $e(N) = \begin{cases} 6, & \text{if } N = 1, \\ 2, & \text{if } N = 2, \\ 1, & \text{if } N \ge 3. \end{cases}$.

Theorem 5.4.4. *[Cf. [Hid81a] Theorem 6.2] $c(f)$ is an integer number unless the following conditions are satisfied:*

1. *k is odd.*
2. *r is odd.*
3. *f is the Mellin transform of a Hecke L-function with a primitive Hecke character λ of an imaginary quadratic field M such that $\lambda(\overline{x}) = \overline{\lambda(x)}$ for all M_A^*.*

If these conditions are satisfied then the nebentype of f is the quadratic residue corresponding to M and $c(f)/\sqrt{d} \in \mathbb{Z}$, where $M = \mathbb{Q}(\sqrt{-d})$.

Corollary 5.4.5. *Let $Z(s,f) = \prod_\sigma L(s, f^\sigma, \chi^\sigma)$. Then,*

$$c(f) = \frac{(\varepsilon(N)(k-1)!NM\varphi(N/M))^r}{2^{r(k+1)}u(f)\pi^{r(k+2)}} Z(k,f),$$

where M is the conductor of χ and $\varepsilon(N) = \begin{cases} 12, & \text{if } N = 1, \\ 4, & \text{if } N = 2, \\ 1, & \text{if } N \geq 3. \end{cases}$.

Proof. The proof of this corollary is just an application of the last two theorems and the formula involving the L-function and the Petersson inner product of the above section. □

5.5 Main Theorem

Now we arrive to the most important Theorem of all this section which gives an insight of what information the L function gives of the behaviour of the function f reduced modulo primes. The result is very interesting in itself because proves the existence of congruences and gives the primes for which these congruences occur. However, as with many other results it proves only existence and gives no method to find g.

Theorem 5.5.1. *Let f be a newform of conductor M and weight $k \geq 2$. Take*

$$C(f) = \begin{cases} \frac{c(f)}{\sqrt{d}}, & \text{if the conditions in 5.4.4 are satisfied} \\ c(f), & \text{otherwise.} \end{cases}$$

Let p be a prime factor of $C(f)$ such that $p > k-2$ and prime to $e_N N$ if $k > 2$, where

$$e_N = \begin{cases} 6, & \text{if } N = 1 \\ 3, & \text{if } N = 2 \\ 1, & \text{if } N \geq 3. \end{cases}$$

Then, there exist a normalised Hecke form g of $\mathcal{S}_k(\Gamma_1(N))$ and a prime $\mathfrak{p}$ dividing p in $\overline{\mathbb{Q}}$ such that

1. *g is not conjugate to f by the action of $G_\mathbb{Q}$.*

2. *$g \equiv f \pmod{\mathfrak{p}}$.*

Proof. Keep the notation of the last section. We have $e(N)^{2r}c(f)^2 = [L_f^* : L_f]$. Let $\overline{M_f}, \overline{L_f}, \overline{L_f^*}$ be the closures in $L_f \otimes_{\mathbb{Z}} \mathbb{Q}_p$ of M_f, L_f, L_f^* and in the same way define $\overline{M_Y}, \overline{L_Y}$. We have $\overline{M_f} = \overline{L_Y}^*$ (cf. [Hid81a] Theorem 3.2), we have that p divides $[M_f : L_f]$. The projection maps of $V(N : R)$ onto $W_f(\mathbb{R}), Y$ induce isomorphisms

$$r_f : L/(L_f \oplus L_Y) \longrightarrow M_f/L_f$$

and

$$r_Y : L/(L_f \oplus L_Y) \longrightarrow M_Y/L_Y$$

with $L = V(N : \mathbb{Z})$. Since f is a newform, these modules have a canonical action of the Hecke algebra and, moreover, this action commutes with r_f, r_Y. Denote by R_f, R_Y the algebras of restricting the Hecke algebra to $W_f(\mathbb{R}), Y$ respectively over $\mathbb{Z}$. This restriction defines a surjective map from the Hecke algebra onto R_f and R_Y. Since p divides$[M_f : L_f]$ there is a maximal ideal $\mathfrak{p}_f$ of R_f of residue characteristic p containing the anihilator of M_f/L_f in R_f, i.e. $\mathfrak{p}_f$ is in the support of M_f/L_f. Let $\mathfrak{p} = \varphi_f^{-1}(\mathfrak{p}_f)$, $\mathfrak{p}_Y = \varphi(\mathfrak{p})$ where φ_f is the ring homomorphism from the Hecke algebra R to R_f.

Then $\mathfrak{p}_Y, \mathfrak{p}$ are non-trivial maximal ideals containing the anihilators of $L/(L_f \oplus L_Y)$ and M_Y/L_Y respectively. We can identify R_f as a ring of integers of K, not necessarily the maximal one, then $\mathfrak{p}_f$ is induced by a prime ideal $\mathfrak{P}$ of $\overline{Q}$. Take $\mathfrak{o} = R_f/|mathfrakp_f$ and identify $\mathfrak{o}$ with $R/\mathfrak{p}$ and $R_Y/\mathfrak{p}_Y$ by the projection φ_f and φ_Y. Then

$$(M_f/L_f) \otimes_{R_f} \mathfrak{o} \simeq (M_Y/L_Y) \otimes_{R_f} \mathfrak{o}$$

as modules over R. On $(M_f/L_f) \otimes_{R_f} \mathfrak{o}$ T_n acts as scalar multiplication of $a_n(f)$ modulo $\mathfrak{P}$ and so does on $(M_Y/L_Y) \otimes_{R_f} \mathfrak{o}$. This representation of R_Y can be lifted to a one dimensional subrepresentation ρ of R_Y in $M_Y \otimes_{\mathbb{Z}} \overline{\mathbb{Q}}$ so that $\rho(T_n) \equiv a_n \pmod{\mathfrak{P}}$, because R_Y is commutative and $R_Y \otimes_{\mathbb{Z}} \mathbb{Q}$ acts on $M_Y \otimes_{\mathbb{Z}} \mathbb{Q}$. Since $M_Y \otimes_{\mathbb{Z}} \mathbb{C} \simeq Y^2$ as $R_f \otimes_{\mathbb{Z}} \mathbb{C}$-modules one can find $g \in Y$ with the desired properties.

□

Chapter 6

Congruences decomposing the space of modular forms

The main aim of this section is to describe the theory of congruences between modular forms that arises from considering decomposition of the space of cusps. This will be of special interest when we treat the cases of old forms and new forms.

6.1 Decomposing the space of modular forms

In this section we will follow the approach of K. Ribet in [Rib83], a nice introduction is given by E. Ghate in [Gha02].

Through this section we will denote by K a field and $\mathcal{O}$ its ring of integers. $M = S(\mathcal{O})$ will denote the space of cusps forms whose coefficients lie in $\mathcal{O}$. In the same way, $S(K)$ will denote the space of cusps whose coefficients lie in K.

We will suppose that $S(K)$ can be decomposed as the direct sum of two spaces

$$S(K) = X \oplus Y$$

In that way, we can define $M_X = M \cap X$ and $M_Y = M \cap Y$, or, using $\pi_X : S(K) \longrightarrow X$ and $\pi_Y : S(K) \longrightarrow Y$ the projection maps and then $M^X = \pi_X(M)$, $M^Y = \pi_Y(M)$. So, we have

$$M_X \oplus M_Y \subseteq M \subseteq M^X \oplus M^Y.$$

Definition 6.1.1. *We define the* ***congruence module***

$$C(M) = \frac{M^X \oplus M^Y}{M} \simeq \frac{M}{M_X \oplus M_Y}.$$

Lemma 6.1.2. *Let $\mathfrak{p}$ be a prime in $\mathcal{O}$, then there are $f \in M_X$ and $f \in M_Y$ such that $f \equiv g \pmod{\mathfrak{p}}$ if and only if $\mathfrak{p} \in \mathrm{Supp}(C(M))$.*

The proof of this fact is very easy putting together the definitions of the support and the definition of $C(M)$.

A useful tool for the study of congruences is, as we have seen several times, the Hecke operators. In this section we will write

$$\mathbb{T} \subseteq \mathrm{End}_{\mathcal{O}}(S(K))$$

the algebra generated by all the Hecke operators. From the inclusion we deduce that its dimension is finite and preserves M.

Alternatively,

Proposition 6.1.3. *If p divides $[M^X \oplus M^Y : M]$ there are $f \in M^X$ and $g \in M^Y$ not divisible by p such that*

$$f - g \in pM$$

. This is the same as

$$f \equiv g \pmod{p}.$$

The proof is immediate from the definitions and hence it gives much information. For instance, there is always a newform and an oldform congruents (if there are oldforms) modulo some prime.

Lemma 6.1.4. $M \simeq \mathrm{Hom}_{\mathcal{O}}(\mathbb{T}, \mathcal{O})$.

Proof. Define the following pairing

$$\begin{array}{ccc} \mathbb{T} \times M & \longrightarrow & \mathcal{O} \\ (T, f) & \longmapsto & a_1(Tf). \end{array}$$

It is $\mathcal{O}$-bilinear and induces two maps

$$\begin{array}{ccccc} M & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\mathbb{T}, \mathcal{O}) & & \\ f & \longmapsto & \varphi_f : \mathbb{T} & \longrightarrow & \mathcal{O} \\ & & T & \longmapsto & a_1(Tf) \end{array}$$

and

$$\begin{array}{ccccc} \mathbb{T} & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(M, \mathcal{O}) & & \\ T & \longmapsto & \psi_f : M & \longrightarrow & \mathcal{O} \\ & & f & \longmapsto & a_1(Tf). \end{array}$$

Both maps are homomorphisms since the original pairing was bilinear and our maps are just fixing one of the components.

From second section, we can assume that M is generated by a basis of Hecke forms, hence, $a_n(g) = a_1(T_n g)$ for any $g \in M$. Then, if we fix f and the first of our maps is the zero homomorphism we have that, for all n

$$a_n(f) = a_1(T_n f) = (T_n, f) = 0$$

so $f = 0$. And the map is injective.

Assume now that we fix T and move f, then

$$0 = (T, f) = a_n(Tf) = a_1(T_nTf) = a_1(TT_nf) = (T, T_nf),$$

since the action of T_n permutes the elements of the basis and the action of T is described by its image of the elements of a basis it follows that $T = 0$. So the second map is injective.

Note that the surjectivity is trivial if we make extension by scalars to K. Assume F is a linear form on $\mathbb{T}$, so we can think it as a K-linear form in $\mathbb{T} \otimes K$, hence there is an element $f \in S(K)$ such that $F(T) = (T, f)$ for all $T \in \mathbb{T} \otimes K$. Taking $T = T_n$ and $F(T_n) = (T_n, f) = a_n(f) \in \mathcal{O}$. Which means $f \in M$. This proves the surjectivity of the first map and proves the lemma. □

Lemma 6.1.5. *Assume that our decomposition is stable under the action of the Hecke operators. Taking $\mathbb{T}^X = \mathbb{T}|_X$ and $\mathbb{T}^Y = \mathbb{T}|_Y$, there is a natural inclusion*

$$\mathbb{T} \hookrightarrow \mathbb{T}^X \oplus \mathbb{T}^Y.$$

Then, $[\mathbb{T}^X \oplus \mathbb{T}^Y : \mathbb{T}]$ is finite.

This is a consequence of the above lemma.

Definition 6.1.6. *We define the **congruence module of the Hecke algebra** with respect to the decomposition of $S(K)$ to be*

$$C(\mathbb{T}) = \frac{\mathbb{T}^X \oplus \mathbb{T}^Y}{\mathbb{T}}.$$

We call it the congruence module for the following reason. Let $\mathfrak{m}$ be a maximal ideal of $\mathbb{T}$ and $\mathfrak{m}_X, \mathfrak{m}_Y$ the respective images in $\mathbb{T}^X, \mathbb{T}^Y$. Then, we have the commutative diagram

$$\begin{array}{ccccc} \mathbb{T}^X & \twoheadleftarrow & \mathbb{T} & \twoheadrightarrow & \mathbb{T}^Y \\ \downarrow & & \downarrow & & \downarrow \\ \mathbb{T}^X/\mathfrak{m}_X & \xleftarrow{\sim} & \mathbb{T}/\mathfrak{m} & \xrightarrow{\sim} & \mathbb{T}^Y/\mathfrak{m}_Y \end{array}$$

Choose minimal prime ideal $\mathfrak{q}_X \subseteq \mathfrak{m}_X$ and $\mathfrak{q}_Y \subseteq \mathfrak{m}_Y$ and let $\mathfrak{p}_X, \mathfrak{p}_Y$ denote their respective pre-images under the maps $\mathbb{T} \twoheadrightarrow \mathbb{T}^X, \mathbb{T}^Y$. So then, we can define $\mathbb{T} \twoheadrightarrow \mathbb{T}/\mathfrak{p}_X$ and $\mathbb{T} \twoheadrightarrow \mathbb{T}/\mathfrak{p}_Y$, which are the same modulo $\mathfrak{m}$. Assume now that we can embed $\mathbb{T}/\mathfrak{p}_X$ and $\mathbb{T}/\mathfrak{p}_Y$ in $\mathcal{O}$ and let $\mathfrak{P}$ denote the maximal ideal of $\mathcal{O}$ corresponding to $\mathfrak{m}$.

Lemma 6.1.4 tells us that from any algebra homomorphism from $\mathbb{T}$ to $\mathcal{O}$ arises an element in M, and since our maps are algebra homomorphisms the element we get is a normalised Hecke form, so in particular we get two cusps that modulo $\mathfrak{P}$ are the same. The goal is to study whether these two cusps are different or they are the same.

Definition 6.1.7. *We will call* ***prime of fusion*** *any* $\mathfrak{p}$ *maximal ideal in* $\mathrm{Supp}C(\mathbb{T})$.

Lemma 6.1.8. *We have* $\mathrm{Supp}(C(\mathbb{T})) = \mathrm{Supp}(C(M))$.

Proof. This follows from the general fact that

$$\mathrm{ann}\left(\frac{\mathbb{T}^X \oplus \mathbb{T}^Y}{\mathbb{T}}\right) = \mathrm{ann}\left(\frac{M^X \oplus M^Y}{M}\right).$$

□

Proposition 6.1.9. $\mathfrak{p} \in \mathrm{Supp}(C(M))$ *if, and only if, there is* $f \in M_X$ *and* $g \in M_Y$ *such that*

$$f \equiv g \pmod{p}$$

where p *is the residue characteristic of* $\mathfrak{p}$.

Proof. By the above results we can choose $h \in \frac{M}{M_X \oplus M_Y}$ with order p. So $\overline{ph} = 0$ then we have $ph \in M_X \oplus M_Y$, i.e., there are $f \in M_X$ and $g \in M_Y$ such that $ph = f - g$, hence

$$f \equiv g \pmod{p}.$$

The converse is easy and follows the same scheme. □

This proposition explains why we call $C(M)$ the congruence module.

Definition 6.1.10. *Let* $\mathfrak{o}_M = \{T \in \mathbb{T} : \ TM \subseteq M\}$.

Corollary 6.1.11. *We have proved hence that the* $\mathfrak{o}_M = \mathbb{T}$.

Proof. This is a consequence of the Lemma 6.1.4. □

Definition 6.1.12. *Let* N *be the level of* M*, we say that a prime number* p *is* ***large*** *if* $(N, p) = 1$*. Otherwise we say that* p *is* ***small****.*

If we make the same construction as $\mathfrak{o}_M$ but now, instead of using M using $\mathfrak{o}_\Lambda$, where Λ is the natural lattice in $S(K)$. We would like to study the primes dividing $(\mathbb{T} : \mathfrak{o}_\Lambda)$.

Proposition 6.1.13. $(\mathbb{T} : \mathfrak{o}_\Lambda)$ *is divisible only by small primes.*

Proof. For the complete proof cf. [Rib83].

We have to differenciate the case when the weight is $k = 2$ and when $k \geq 3$.

For $k = 2$ the idea is to fix a large prime p and to prove that $\mathbb{T}/p\mathbb{T}$ acts faithfully on $\Lambda/p\Lambda$, i.e., if T annihilates $\Lambda/p\Lambda$ must annihilate too M/pM.

Let Λ_k be the lattice spanned over $\mathcal{O}$ by a Hecke basis of $S_k(K)$ and let $d(k) = \dim S_k(K)$ define

$$V_k = \Lambda_k / p\Lambda_k,$$

and

$$d = d(3) \oplus \cdots \oplus d(p).$$

We have now that

$$2d = \dim_{\mathbb{F}_p}(V).$$

Let $R_k = \mathbb{T}|_{V_k}$, and

$$R = R_3 \oplus \cdots \oplus R_p.$$

Assume for the rest of the proof that $\dim_{\mathbb{F}_p}(R) \geq d$. What we want to prove is equivalent to

$$i_k : \mathbb{T}_k / p\mathbb{T}_k \longrightarrow \mathrm{End}(V_k)$$

is injective for any k. By Lemma 6.1.4

$$\dim \mathbb{T}_k / p\mathbb{T}_k = d(k).$$

Hence,

$$\dim \mathrm{Im}(i_k) \leq d(k),$$

and our goal is to prove the equality. By definition we know that $\mathrm{Im}(i_k) = R_k$, we have the inclusion

$$R \subseteq \bigoplus_{k=3}^{p} R_k,$$

this implies

$$\dim_{\mathbb{F}_p}(R) \leq \sum_{k=3}^{p} \dim_{\mathbb{F}_p} R_k \leq \sum_{k=3}^{p} d(k) = d.$$

So by the assumption that $\dim_{\mathbb{F}_p}(R) \geq d$ we deduce that we have indeed an equality and this proves the proposition. □

So our proof is held by the following theorem that we will not prove.

Theorem 6.1.14 (Cf. [Rib83] sections 3 and 4)**.** *With the same notation,*

$$\dim_{\mathbb{F}_p}(R) \geq d.$$

6.2 Discriminants

As in the preceding chapter we can associate to a lattice L in a space V a number called the **discriminant**. This notion, as in the preceding chapter is enough to gave some conditions on the existence of congruences between modular forms.

Definition 6.2.1. *Let V be a finite dimensional $\mathbb{Q}$-vector space and*

$$\omega : V \times V \longrightarrow \mathbb{Q}$$

*a $\mathbb{Q}$-bilinear form non-degenerate. Let Λ be a lattice in V such that $\omega(\Lambda, \Lambda) \subseteq \mathbb{Z}$. We define the **discriminant** of Λ by*

$$d(\Lambda) = |\det(\omega(e_i, e_j))|,$$

where $\{e_1, \ldots, e_r\}$ is a basis of Λ.

The discriminant is well defined, indeed, if $\{f_1, \ldots, f_r\}$ is another basis, we have that there is a matrix M with integral coefficients (because $\omega(\Lambda, \Lambda) \subseteq \mathbb{Z}$) which is invertible. Then

$$|\det(\omega(e_i, e_j))| = |\det M|^2 |\det(\omega(f_i, f_j))| = |\det(\omega(f_i, f_j))|.$$

Proposition 6.2.2 (Cf. [Ser79] Chapter 3, Section 2, Proposition 5)**.** *Let Λ_1, Λ_2 be two lattices in V. Assume there is an exact sequence*

$$0 \longrightarrow \Lambda_1 \longrightarrow \Lambda_2 \longrightarrow \Lambda_2/\Lambda_1 \longrightarrow 0,$$

then $[\Lambda_2 : \Lambda_1]$ is finite.

Let us assume that V is a space of cusp forms of some weight k, level N and nebentypus χ not containing old forms and such that there is a basis B whose elements have integral coefficients. For any $f \in B$ let us define $X_f = \{f^\sigma : \sigma \in G_{\mathbb{Q}}\}$. So we have

$$S(\mathbb{Q}) = \bigoplus_{f \in B} X_f.$$

Similarly, if we denote by $\mathbb{T}_f$ the subalgebra of $\mathrm{End}_{\mathbb{Z}}(X_f)$ generated by the Hecke operators we then have that $\mathbb{T}, \oplus_{f \in B} \mathbb{T}_f$ are lattices in $\mathbb{T} \otimes \mathbb{Q}$ related by the sequence

$$0 \longrightarrow \mathbb{T} \longrightarrow \oplus_{f \in B} \mathbb{T}_f \longrightarrow \left(\oplus_{f \in B} \mathbb{T}_f\right)/\mathbb{T} \longrightarrow 0.$$

Moreover, we can endow a natural bilinear form in $\mathbb{T} \otimes \mathbb{Q}$ by $\omega(A, B) = \mathrm{Tr}(AB)$. It takes values in $\mathbb{Z}$ for $\mathbb{T}$ and $\mathbb{T}_f$.

By the last proposition

$$d(\mathbb{T}) = \left|\left(\oplus_{f \in B} \mathbb{T}_f\right)/\mathbb{T}\right|^2 \prod_{f \in B} d(\mathbb{T}_f).$$

If $\mathcal{O}$ is the ring of integers generated by the coefficients f, there is an isomorphism

$$\begin{array}{ccc} \mathbb{T}_f & \longrightarrow & \mathcal{O} \\ T_n & \longmapsto & a_n(T_n f). \end{array}$$

Then $d(\mathbb{T}_f) = d(\mathcal{O})$.

Proposition 6.2.3. *Let f be a cusp form of weight k, level N and nebentypus χ, not an old form. Let K be the field spanned by the Fourier coefficients of f and let $\mathcal{O}$ be the ring of integers of K. Assume the Fourier coefficients of f are in $\mathcal{O}$ and that K is Galois. Then $p|d(\mathcal{O})$ if, and only if, there is a prime $\mathfrak{p}$ of $\mathcal{O}$ dividing p and a non trivial element $\sigma \in \mathrm{Gal}(K|\mathbb{Q})$ such that*

$$f^\sigma \equiv f \pmod{\mathfrak{p}}.$$

Proof. Assume $p|d(\mathcal{O})$ and let $\mathfrak{p} \subseteq \mathcal{O}$ be a prime above p. Let $I_\mathfrak{p}$ be the inertia group at $\mathfrak{p}$. Since p ramifies there is $\sigma \in I_\mathfrak{p}$ non-trivial. Since

$$\sigma(x) \equiv x \pmod{\mathfrak{p}},$$

for all $x \in \mathcal{O}$ it is true for $a_n(f)$, i.e.

$$\sigma(a_n(f)) = a_n(f^\sigma) \equiv a_n(f) \pmod{\mathfrak{p}}.$$

Hence,

$$f^\sigma \equiv f \pmod{\mathfrak{p}}.$$

Conversely, if $\mathfrak{p} \subseteq \mathcal{O}$ and for a non-trivial $\sigma \in \mathrm{Gal}(K|\mathbb{Q}) \setminus \{\mathrm{Id}\}$ we have that

$$f^\sigma \equiv f \pmod{\mathfrak{p}},$$

this implies that

$$\sigma(x) \equiv x \pmod{\mathfrak{p}},$$

for all $x \in \mathcal{O}$ and therefore $\sigma \in I_\mathfrak{p}$ which implies that p ramifies and that $p|d(\mathcal{O})$. □

Bibliography

[BC67] Z.I. Borevitch and I.R. Chafarevitch. *Théorie des nombres*. Gauthier-Villars, 1967.

[Coh07] H. Cohen. *Number Theory Volume I: Tools and Diophantine Equations*. Graduate Texts in Mathematics. Springer-Verlag, 2007.

[Cox89] David A. Cox. *Primes of the Form $x^2 + ny^2$*. A Wiley-Interscience Series of Texts, Monographs, and Tracts. John Wiley and sons INC, 1989.

[Del69] P. Deligne. Formes modulaires et représentations ℓ-adiques. *Séminaire Borubaki*, 355:139–172, 1969.

[DF14] N. Dummigam and D. Fretwell. Ramanujan style congruences of local origin. *Journal of Number Theory*, 143:248–261, 2014.

[DG96] B. Datskovsky and P. Guerzhoy. On Ramanujan congruences for modular forms of integral and half-integral weights. *Proceedings of the American Mathematical Society*, 124(8):2283–2291, 1996.

[DI95] F. Diamond and J. Im. Modular forms and modular curves. In *Canadian Mathematical Society Conference Proceedings*, volume 17, pages 39–133, 1995.

[DS05] F. Diamond and J. Shurman. *A First Course in Modular Forms*. Graduate Texts in Mathematics. Springer-Verlag, 2005.

[FK90] R. Fricke F. Klein. Voresungen über die Theorie der elliptischen Functionen. *Teubnen*, 1890.

[Gha02] E. Ghate. An introduction to congruences between modular forms. In S. A. Katre S. D. Adhikari and B. Ramakrishnan, editors, *Current Trends in Number Theory*, pages 39–58, Gurgaon, 2002. Hindustan Book Agency.

[Hec24] E. Hecke. Analytische Funktionen und algebraische Zahlen, II. *Hamb. Math. Abh.*, (3):213–236, 1924.

[Hid81a] H. Hida. Congruence of cusp forms and special values of their zeta functions. *Invent. Math.*, 63:225–261, 1981.

[Hid81b] H. Hida. On congruence divisors of cusp forms as factors of the special values of their zeta functions. *Invent. Math.*, 64:221–262, 1981.

[IC10] J. Rasmussen I. Chen, I. Kiming. On congruences mod $\mathfrak{p}^m$ between eigenforms and their attached galois representations. *Journal of Number Theory*, (130):608–619, 2010.

[KD76] T. Miyake K. Doi. *Automorphic forms and number theory (in Japanese).* Kinokuniya Shoten, 1976.

[Leh43] D. H. Lehmer. Ramanujan's function $\tau(n)$. *Duke Math. J.*, (10):483–492, 1943.

[Miy71] T. Miyake. On automorphic forms on gl_2 and hecke operators. *Ann. of Math.*, 94:174–189, 1971.

[Miy06] T. Miyake. *Modular Forms.* Springer monographs in mathematics. Springer, 2006.

[Mur97] M. Ram Murty. Congruences between modular forms. *London Mathematical Society Lecture Notes*, 247:309–320, 1997.

[Ono94] K. Ono. Congruences on the fourier coefficients of modular forms on $\Gamma_0(N)$. *Contemporary Mathematics*, 166:93–105, 1994.

[Rib83] K. Ribet. Congruence relations between modular forms. In *Proceedings of the Internationl Congress of Mathematicians*, volume II, pages 503–514, 1983.

[SD73] H.P.F. Swinnerton-Dyer. On ℓ-adic representations and congruences for coefficients of modular forms. *Springer Lecture Notes in Math.*, (350):1–55, 1973.

[Ser79] J.-P. Serre. *Local fields.* Graduate Texts in Mathematics. Springer, 1979.

[Ser89] J.-P. Serre. *Abelian ℓ-adic Representations and Elliptic Curves.* The advanced book program. Addison-Wesley Publishing Company, 1989.

[Shi71] G. Shimura. *Introduction to the Arithmetic Theory of Automorphic Functions.* Iwanami Shoten and Princeton Univ. Press, 1971.

[Shi73] G. Shimura. On the factors of the jacobian variety of a modular function field. *J. Math. Soc.*, 25:523–544, 1973.

[Shi75] G. Shimura. On the holomorphy of certain dirichlet series. *Proc. London Math. Soc.*, 31:79–98, 1975.

[Shi76] G. Shimura. The special values of the zeta functions associated with cusp forms. *Comm. pure appl. Math.*, 29:783–804, 1976.

[Stu84] J. Sturm. On the Congruence of Modular Forms. *Springer Lecture notes in Mathematics*, 1240, 1984.